Stochastic differential equations with multiple invariant measures and related problems

Paul Dobson

Submitted for the degree of
Doctor of Philosophy

Heriot-Watt University

Department of Mathematics,
School of Mathematical and Computer Sciences.

August 20, 2019

Abstract

We study problems related to SDEs which admit multiple invariant measures. The main problem we address is determining the long time behaviour of a large class of diffusion processes on $\mathbb{R}^N$, generated by second order differential operators of (possibly) degenerate type. The operators that we consider *need not* satisfy the Hörmander condition and *need not* admit a unique invariant measure. Instead, we consider the so-called UFG condition, introduced by Hermann, Lobry and Sussmann in the context of geometric control theory and later by Kusuoka and Stroock, this time with probabilistic motivations. In this thesis we will demonstrate the importance of UFG diffusions in several respects: We show that UFG processes constitute a family of SDEs which exhibit multiple invariant measures and for which one is able to describe a systematic procedure to determine the basin of attraction of each invariant measure (equilibrium state). We show that our results and techniques, which we devised for UFG processes, can be applied to the study of the long-time behaviour of non-autonomous hypoelliptic SDEs. We prove that there exists a change of coordinates such that every UFG diffusion can be, at least locally, represented as a system consisting of an SDE coupled with ODE, where the ODE evolves independently of the SDE part of the dynamics. As a result, UFG diffusions are inherently "less smooth" than hypoelliptic SDEs; more precisely, we prove that UFG processes do not admit a density with respect to Lebesgue measure on the entire space, but only on suitable time-evolving submanifolds, which we describe. We introduce a novel pathwise approach to obtain (long-time) derivative estimates for Markov semigroups. The content of this thesis has resulted in two long papers [1] and [2], both submitted for publication.

Acknowledgements

I would like to thank my supervisor Michela Ottobre for her patience and guidance throughout my PhD. I also wish to thank both Dan Crisan and Tom Cass for their help and the many useful conversations we have had. To the many PhD students in Edinburgh who have been willing to listen to me while I get my thoughts in order, offer support and suggestions and mostly making Edinburgh a pleasant place to live and work. The community in Edinburgh has made the last four years a very enjoyable time; especially the reading groups I have taken part in and the probability working seminar for being a welcoming group to study and discuss ideas with. I wouldn't have been able to have this opportunity if it wasn't for the funding provided by The Maxwell Institute Graduate School in Analysis and its Applications, a Centre for Doctoral Training funded by the UK Engineering and Physical Sciences Research Council (grant EP/L016508/01), the Scottish Funding Council, Heriot-Watt University and the University of Edinburgh. All of MIGSAA but in particular my cohort has been a supporting a friendly group to be a part of and made arriving in Edinburgh and life in the city a fun and memorable time. Finally, I would like to thank my family for supporting and encouraging me.

ACADEMIC REGISTRY

Research Thesis Submission

Please note this form should be bound into the submitted thesis.

Name:	Paul Dobson		
School:	MACS		
Version: *(i.e. First, Resubmission, Final)*	Final	Degree Sought:	PhD

Declaration

In accordance with the appropriate regulations I hereby submit my thesis and I declare that:

1. The thesis embodies the results of my own work and has been composed by myself
2. Where appropriate, I have made acknowledgement of the work of others
3. Where the thesis contains published outputs under Regulation 6 (9.1.2) these are accompanied by a critical review which accurately describes my contribution to the research and, for multi-author outputs, a signed declaration indicating the contribution of each author (complete Inclusion of Published Works Form – see below)
4. The thesis is the correct version for submission and is the same version as any electronic versions submitted*.
5. My thesis for the award referred to, deposited in the Heriot-Watt University Library, should be made available for loan or photocopying and be available via the Institutional Repository, subject to such conditions as the Librarian may require
6. I understand that as a student of the University I am required to abide by the Regulations of the University and to conform to its discipline.
7. Inclusion of published outputs under Regulation 6 (9.1.2) shall not constitute plagiarism.
8. I confirm that the thesis has been verified against plagiarism via an approved plagiarism detection application e.g. Turnitin.

* *Please note that it is the responsibility of the candidate to ensure that the correct version of the thesis is submitted.*

Signature of Candidate:		Date:	

Submission

Submitted By *(name in capitals)*:	PAUL DOBSON
Signature of Individual Submitting:	
Date Submitted:	

For Completion in the Student Service Centre (SSC)

Received in the SSC by *(name in capitals)*:			
Method of Submission *(Handed in to SSC; posted through internal/external mail)*:			
*E-thesis Submitted (**mandatory for final theses)***			
Signature:		Date:	

Contents

Chapter 1

Introduction

1.1 Overview

The study of diffusion processes of hypoelliptic type has by now produced a fully-fledged theory, involving several branches of mathematics: stochastic analysis, analysis of differential operators, Riemannian (or sub-Riemannian) geometry and control theory. One of the key steps in the development of such a theory has been the seminal paper of Hörmander [3] and a large body of work has been dedicated for over forty years to the study of diffusion processes under the Hörmander Condition (HC) (in one of its many forms), which is a sufficient condition for hypoellipticity. In particular, the *ergodic theory* for Hörmander type SDEs is well developed, see [4–11] and references therein. The notions of ellipticity, hypoellipticity and the Hörmander condition will be rigorously recalled in Chapter 2 and Appendix A.1.1. As for the notion of ergodicity, we clarify that throughout this thesis we define a process to be *ergodic* if it admits a unique *invariant measure* (also referred to in this thesis as a stationary state or equilibrium measure).

To the best of our knowledge, this thesis is the first attempt to build a consistent framework for the study of the long time asymptotics of solutions of SDEs which admit several invariant measures (i.e. with our nomenclature, SDEs which are not ergodic); while there are several examples of Hörmander type SDEs which exhibit more than one invariant measure (we will further elaborate on this point in Section 3.1), we found that requiring the validity of the (parabolic) Hörmander condition imposes a lot of structure on the geometrical properties of the SDE. Such a structure is not natural if one is interested in a non-ergodic setting. For this reason, the SDEs that we consider here *need not* satisfy the (parabolic) Hörmander condition (and need not be hypoelliptic).

To explain in more detail the content of this thesis let us first introduce some standard notation. We will consider stochastic differential equations (SDEs) in $\mathbb{R}^N$ of the form

$$X_t^{(x)} = x + \int_0^t V_0(X_s^{(x)})ds + \sqrt{2}\sum_{i=1}^{d}\int_0^t V_i(X_s^{(x)}) \circ dB^i(s), \quad X_0^{(x)} = x, \tag{1.1}$$

where $V_0, \dots, V_d$ are smooth vector fields on $\mathbb{R}^N$, $\circ$ denotes Stratonovich integration and $B^1(t), \dots, B^d(t)$ are one dimensional independent standard Brownian motions. The superscript (x) is to emphasise the dependence of the solution on the initial datum x. The Markov semigroup $\{\mathcal{P}_t\}_{t\geq 0}$ associated with the SDE (1.1) is defined on the set $C_b(\mathbb{R}^N)$ of continuous and bounded functions as

$$\mathcal{P}_t : C_b(\mathbb{R}^N) \to C_b(\mathbb{R}^N), \quad (\mathcal{P}_t f)(x) := \mathbb{E}\left[f(X_t^{(x)})\right]. \tag{1.2}$$

We recall that, given a vector field $V : \mathbb{R}^N \to \mathbb{R}^N$, we can interpret V both as a vector-valued function on $\mathbb{R}^N$ and as a first order differential operator on $\mathbb{R}^N$:

$$V = (V^1(x), V^2(x), \dots, V^N(x)) \quad \text{or} \quad V = \sum_{j=1}^{N} V^j(x)\partial_j, \quad x \in \mathbb{R}^N, \partial_j = \partial_{x^j}. \tag{1.3}$$

With this notation, the Kolmogorov operator associated with the semigroup $\mathcal{P}_t$ is the second order differential operator given on smooth functions by

$$\mathcal{L} = V_0 + \sum_{i=1}^{d} V_i^2, \tag{1.4}$$

and a probability measure μ on $\mathbb{R}^N$ is an *invariant measure* for the semigroup $\mathcal{P}_t$ if for all $f \in C_b(\mathbb{R}^N)$ we have

$$\int_{\mathbb{R}^N} \mathcal{P}_t f(y)\mu(dy) = \int_{\mathbb{R}^N} f(y)\mu(dy).$$

As already mentioned, we will not assume that the vector fields $V_0, \dots, V_d$ satisfy the Hörmander condition. We will instead work in the setting in which the vector fields $V_0, \dots, V_d$ satisfy a weaker condition, the so-called *UFG condition*. The acronym UFG stands for *Uniformly Finitely Generated*. Before moving on to explaining the nature of the UFG condition, we would like to further remark on the lack of ergodicity.

While a vast theory exists regarding the study of ergodic processes, both in finite and infinite dimensions, see [8–10, 12–17], the development of a general framework to understand problems with multiple equilibria is at a very early stage. It is well known that ergodic processes will, under appropriate conditions, converge to their unique equilibrium measure irrespective of the initial configuration, i.e. they will tend to lose memory of the initial datum. Indeed, typical methods in ergodic theory are aimed at proving results of the type

$$\mathcal{P}_t f(x) \to \int_{\mathbb{R}^N} f(y)\mu(dy), \qquad \text{as } t \to \infty.$$

Note that the initial datum x of the SDE does appear on the left hand side but it is not present on the right hand side. Clearly this type of result cannot hold, in general, for more complicated systems. When the invariant measure is not unique it is typically extremely difficult to determine the basin of attraction of each equilibrium measure and we are indeed not aware of any general criteria developed to this effect. To be more precise, one can ask one of the two (complementary) questions: given an initial datum for the SDE, which equilibrium measure will the process converge to? Conversely, given an equilibrium measure μ, what is the basin of attraction of such a measure, i.e. the set of initial data $x \in \mathbb{R}^N$ such that the process $X_t^{(x)}$ converges to μ.

In this thesis we introduce a systematic way to study long-time convergence for a large class of SDEs which will, in general, admit several stationary states. This methodology applies to UFG diffusions and hence, because processes that satisfy the (uniform) parabolic Hörmander condition are UFG processes, our results produce further understanding on non-ergodic Hörmander processes.

We give a precise statement of the UFG condition in Definition 2.2.1 and compare it carefully with the (uniform) Parabolic Hörmander Condition (and with other forms of the Hörmander Condition) in Appendix A.1.2 and in Chapter 2. For the moment let us just informally point out that, while the Parabolic Hörmander Condition (PHC) imposes the following

$$\mathrm{span}\{\mathfrak{L}_j(x) : j \geq 1\} = \mathbb{R}^N \quad \text{for every } x \in \mathbb{R}^N, \tag{\textbf{PHC}}$$

where as customary the hierarchy of operators $\mathfrak{L}_j$ is defined as

$$\mathfrak{L}_1(x) := \{V_1(x), \ldots, V_d(x)\}$$

$$\mathfrak{L}_j(x) := \mathfrak{L}_{j-1}(x) \cup \{[V, V_k], V \in \mathfrak{L}_{j-1}, k \in \{0, \ldots, d\}\}, \qquad \text{for any } j > 1,$$

under the UFG condition the vector space appearing in (**PHC**) is not required to have constant rank; it is only required to be *finitely generated*. In particular, we emphasize that the UFG condition does not impose the vector space in (**PHC**) to be homeomorphic to $\mathbb{R}^N$ for any $x \in \mathbb{R}^N$. Hence, in this sense, the UFG condition is weaker than the (uniform) parabolic Hörmander condition. Consider for example the SDE

$$dX_t = X_t dt + \sqrt{2} X_t \circ dB_t. \tag{1.5}$$

where $\{B_t\}_{t \geq 0}$ is a one-dimensional Brownian motion. As is well known, we can solve (1.5) explicitly to find

$$X_t = X_0 \exp(t + \sqrt{2} B_t). \tag{1.6}$$

This is the simplest example of an SDE which is of UFG type but does not satisfy the parabolic Hörmander condition. In this case, $V_1 = V_0 = x\partial_x$ and the associated Kolmogorov operator is given by

$$\mathcal{L} = V_1 + V_1^2.$$

Clearly here $\mathfrak{L}_1(x) = \{V_1(x)\} = \mathfrak{L}_j(x)$, for every $j \geq 1$ so, because $V_1(x)$ vanishes at $x = 0$, the (**PHC**) is not satisfied by (1.5).

In passing, we note that from (1.6) one can also make the following simple observation: if $X_0 > 0$ then $X_t > 0$ for all $t \geq 0$; similarly, if $X_0 < 0$ then the solution stays negative and if $X_0 = 0$ then $X_t = 0$ for all $t \geq 0$. We will see that this splitting of state space is a general characteristic of UFG processes and an important part of this thesis is devoted to giving a procedure to describe the manifolds in which the space $\mathbb{R}^N$ is naturally partitioned by the process, see Chapter 3.

The UFG condition has been long known by the (geometric) control theory community, although perhaps under other names (see Section 2.2 for a more detailed account on the matter), and it is well-studied in the works of Hermann, Lobry and Sussmann [18–20]. It was then considered by Stroock and Kusuoka in the eighties [21–24], though in a completely different context (which we briefly explain below). The purpose there was to study smoothing properties of the semigroup $\mathcal{P}_t$ under the UFG condition. In this thesis we combine the geometric viewpoint with the functional analytic and probabilistic one to introduce new results on the asymptotic behaviour of UFG diffusions. In broad terms, the three main achievements of this thesis can be described as follows:

i) We study the diffusion process (1.1) in absence of the Hörmander condition. To this end, we establish explicit connections between the geometric theory of finitely generated Lie algebras and the related stochastic dynamics. Because every (uniformly) Hörmander process is a UFG process,[1] our results cover a very large class of SDEs. We will show that, as a byproduct, our approach can be fruitfully employed to study the asymptotic behaviour of non-autonomous hypoelliptic diffusions as well.

ii) We argue that UFG processes constitute a class of SDEs which exhibit, in general, multiple equilibria and for which one is able, given an initial datum, to determine the invariant measure to which the dynamics will converge.

iii) On a technical level, to deal with the presence of many invariant measures, one needs long-time estimates on the space derivatives of the Markov semigroup $\mathcal{P}_t$. Here we develop a novel approach which allows one to obtain sufficient conditions under which the derivatives of the semigroup decay exponentially fast in time. The technique we develop here is a pathwise version of the Bakry-Emery approach, see [25]. In this thesis we use such derivative estimates to understand the long-time behaviour of the semigroup $\mathcal{P}_t$. However we emphasise that such estimates play a

[1] See Section 2.2

fundamental role in [2] as well, where they are instrumental to obtaining sufficient criteria for the Euler method to approximate the SDE (1.1) with a weak error which converges to 0 *uniformly in time*. Another case where estimates of this type play an important role is in [26, 27], the aim of these papers is to understand an interacting particle system in either a large time limit or as the number of particles tends to infinity.

Throughout, we will also present a wealth of examples to which our theory applies. Many of the examples we include are designed to give a simple illustration of the ideas we present. However we will also consider more complex SDEs such as the Stochastic Geodesic Equation (SGE), see Example 3.1.8. The SGE exhibits multiple invariant measures and it has been studied in [28]. When studied on a "case by case" basis, such an SDE is quite challenging to understand. We will show that the long time behaviour of the SGE becomes straightforward to analyse if one makes use of the theory of UFG processes that we develop in this thesis.

The Markov diffusions studied in this manuscript are *linear*, in the sense that their generators (1.4) are linear second order differential operators. As a point of comparison, another class of systems exhibiting multiple equilibria is the class of so-called *collective dynamics*: in this case the system is constituted by a large number of particles or agents that interact with each other. The underlying kinetic-PDEs for this type of models are *non-linear in the sense of McKean* and the existence of multiple stationary states here is due to such a nonlinearity. In our case, the nature of the phenomenon is completely different and in a way simpler, as multiple invariant measures arise as a result of the non-trivial control-theory implied by the UFG condition.

Let us now comment on the implications and significance of the UFG condition first from an analytic perspective and then from a geometric and probabilistic viewpoint.

UFG condition in analysis. As is well known, under the (parabolic) Hörmander condition, the transition probabilities of the semigroup $\mathcal{P}_t$ have a smooth density (we recall why this is the case in Appendix A.1.1); furthermore, $\mathcal{P}_t f$ is differentiable in every direction and $u(t,x) := (\mathcal{P}_t f)(x)$ is a classical solution of the following

Cauchy problem

$$\partial_t u(t,x) = \mathcal{L}u(t,x)$$

$$u(0,x) = f(x).$$

In a series of papers [21–24, 29–31], Kusuoka and Stroock first and Crisan and collaborators later, have analyzed the smoothness properties of diffusion semigroups $\{\mathcal{P}_t\}_{t\geq 0}$ associated with the stochastic dynamics (1.1) when the vector fields $\{V_i, i = 0, 1, ..., d\}$ satisfy the UFG condition. Such works showed that, as opposed to what happens under the PHC, under the UFG condition the semigroup $\mathcal{P}_t$ is no longer differentiable in every direction; in particular it is no longer differentiable in the direction V_0, but it is still differentiable in the direction $\mathcal{V} := \partial_t - V_0$ when viewed as a function $(t,x) \mapsto u(t,x)$ over the product space $(0,\infty) \times \mathbb{R}^N$. This fact has been proved by means of Malliavin calculus in [21–24] and in this thesis we give a geometric and analytic explanation of such a phenomenon. Because of differentiability in the direction $\mathcal{V}$, a rigorous PDE analysis can still be built starting from the stochastic dynamics (1.1). In this case one can indeed prove that for every $f \in C_b$ (continuous and bounded), the function $u(t,x) := (\mathcal{P}_t f)(x)$ is a classical solution[2] of the Cauchy problem

$$\begin{cases} \mathcal{V}u(t,x) &= \sum_{i=1}^d V_i^2 u(t,x) \\ u(0,x) &= f(x). \end{cases} \tag{1.7}$$

A simple example to illustrate differentiability in the direction $\mathcal{V}$ is given by the one-dimensional transport equation, namely

$$\begin{cases} \partial_t u(t,x) &= \partial_x u(t,x), \quad t > 0, x \in \mathbb{R}, \\ u(0,x) &= f(x), \qquad\qquad\qquad x \in \mathbb{R}. \end{cases} \tag{1.8}$$

As is well known the solution of this PDE is given by $u(t,x) = f(x+t)$. With our notation here $N = 1, d = 1$, $V_0 = \partial_x$ and $V_1 = 0$. We have not yet rigorously introduced the UFG condition however the transport equation is an extreme example

[2]The notion of classical solution for the PDE (1.7) and further background material can be found in [32, Appendix A].

of a dynamics satisfying such a condition. Clearly if f is not smooth then u is also not smooth in either the direction ∂_x or ∂_t, however u is constant in the direction $\mathcal{V} := \partial_t - \partial_x$ in particular, it is smooth in the direction $\mathcal{V}$.

UFG condition in geometry and control theory. From a geometrical and control-theoretical point of view, working with the UFG condition will imply dealing with distributions of non-constant rank. If the geometric understanding of the Hörmander condition is rooted in the classic Frobenius Theorem, which deals with distributions of constant rank, the geometry of the UFG condition is described in the works of Hermann, Lobry and Sussmann [18–20]. In these works, the UFG condition was considered for geometric and control theoretical purposes, in particular for the study of reachability (i.e., roughly speaking, to answer questions regarding the set of points that can be reached by the integral curves of given vector fields). In this respect we should stress that the UFG condition is not optimal from a control-theoretical point of view (an optimal condition for reachability has been described by Sussmann [20]). However, it is the closest to being optimal, while still being easy to check in practice.

UFG condition in probability. Finally, by a probabilistic standpoint, it is well known that the Hörmander condition is a sufficient (and almost necessary) condition for the law of the process (1.1) to have a density, see [33–35], and this fact has motivated the large literature on hypoelliptic SDEs. Again, the understanding of this matter relies on Frobenius Theorem, as Hörmander himself noted [3]. In his seminal paper [36], Bismut proved that, when the Hörmander condition (HC) is enforced in place of the PHC,[3] the law of the process no longer admits a density on $\mathbb{R}^N$; however, it admits a density on appropriate time-dependent submanifolds of $\mathbb{R}^N$. In this thesis we prove that a similar statement holds, in more generality, for UFG processes, and in Chapter 6 we explicitly describe the time-dependent manifolds on which the process admits a law.

[3] The difference between the PHC and the HC will be clarified in Chapter 2.

1.2 Main Results

Let us give a rather informal description of the main results of this thesis. Precise notation, assumptions and statements are deferred to the relevant sections.

A *distribution* Δ on $\mathbb{R}^N$ is a map that, to each point $x \in \mathbb{R}^N$, associates a linear subspace of the tangent space $T_x\mathbb{R}^N$. Given a set $\mathcal{D}$ of smooth vector fields on $\mathbb{R}^N$, the distribution generated by $\mathcal{D}$, denoted by $\Delta_{\mathcal{D}}$, is the map $x \mapsto \mathrm{span}\{X(x) : X \in \mathcal{D}\}$. Let us introduce two distributions, $\hat{\Delta}(x)$ and $\hat{\Delta}_0(x)$, that will play a fundamental role in this thesis. To avoid having to set too much notation and nomenclature, we introduce them now informally but we will give precise definitions at the beginning of Section 3.1.[4] The distribution $\hat{\Delta}$ is generated by the vector fields contained in the Lie algebra (**PHC**), i.e. it is the distribution

$$\hat{\Delta}(x) = \bigcup\{\mathfrak{L}_j(x) : j \geq 1\} \tag{1.9}$$

while

$$\hat{\Delta}_0(x) = \mathrm{Lie}\{V_0(x), V_1(x), , \ldots, V_d(x)\} \tag{1.10}$$

$$= \mathrm{span}\{V_0(x)\} + \hat{\Delta}(x)\,. \tag{1.11}$$

Clearly, $\hat{\Delta}(x) \subseteq \hat{\Delta}_0(x)$ for every $x \in \mathbb{R}^N$ and the two distributions coincide at x if and only if $V_0(x)$ is a combination of the vectors contained in $\hat{\Delta}$. More precisely, we decompose the vector V_0 into a component which belongs to $\hat{\Delta}$, $V_0^{(\hat{\Delta})}$, and a component which is orthogonal to $\hat{\Delta}$, $V_0^{(\perp)}$:

$$V_0 = V_0^{(\hat{\Delta})} + V_0^{(\perp)}\,. \tag{1.12}$$

In other words, $V_0^{(\perp)}(x)$ is the orthogonal projection of $V_0(x)$ on the vector space $\hat{\Delta}(x)$, so the two distributions coincide if and only if $V_0^{(\perp)} = 0$. We will see that the vector $V_0^{(\perp)}$ plays an important role for the dynamics and, ultimately, it is the component of V_0 responsible for the lack of smoothness in the direction V_0.

[4] In that section we define them differently, but we then prove that the definition we give there is equivalent to the one we state in (1.9) - (1.10).

Therefore, in a way, the distribution $\hat{\Delta}$ is the one containing all the directions along which the problem (1.7) is smooth. We will come back to this later.

Under the UFG condition the integral manifolds (see Section 2.3 for definition) of $\hat{\Delta}_0$ form a partition of the state space $\mathbb{R}^N$. Let $\mathscr{S}$ be one such manifold.[5] If $X_0 = x \in \mathscr{S}$ then $X_t^{(x)} \in \overline{\mathscr{S}}$ for all $t \geq 0$. That is, if the process starts from one of the manifolds of the partition, then it remains in the closure of such a manifold (just like in (1.5)-(1.6)); we stress that, crucially, the process can hit the boundary $\partial\mathscr{S} := \overline{\mathscr{S}} \setminus \mathscr{S}$ of the manifold $\mathscr{S}$. This is the content of Proposition 3.1.7. Such a statement is obtained by combining the known geometric theory of distributions with non-constant rank and the classical Stroock–Varadhan support theorem. We further prove that if X_t hits the boundary $\partial\mathscr{S}$ of the manifold $\mathscr{S}$, then it never leaves it, see Proposition 3.2.1 and Note 3.2.2. Therefore: i) because the dimension of the boundary $\partial\mathscr{S}$ is smaller than the dimension of $\mathscr{S}$, along the path of $X_t^{(x)}$ the rank of the distribution cannot increase; ii) if the solution of the SDE leaves the manifold $\mathscr{S}$ from where it started, then any invariant measure can only be supported on the boundary $\partial\mathscr{S}$ of such a manifold, see Proposition 3.2.7.

Further understanding of the dynamics relies on the results of Section 3.1.2: in this section we show that, after an appropriate change of coordinates, any N-dimensional SDE of UFG-type can be written, at least locally, as a system of the form

$$dZ_t = U_0(Z_t, \hat{\zeta}_t)dt + \sum_{j=1}^{d} U_j(Z_t, \hat{\zeta}_t) \circ dB_t^j \tag{1.13}$$

$$d\hat{\zeta}_t = \hat{W}_0(\hat{\zeta}_t)dt\,, \tag{1.14}$$

where $\hat{\zeta}_t$ solves an ordinary differential equation (ODE), $\hat{\zeta}_t \in \mathbb{R}^{N-n}$,[6] $Z_t \in \mathbb{R}^n$, $\hat{W}_0 : \mathbb{R}^{N-n} \to \mathbb{R}^{N-n}$ and $U_i : \mathbb{R}^N \to \mathbb{R}^n$ for every $i \in \{0, \ldots, d\}$. Beyond details about the dimensionality of the ODE component, the important thing is that the solution of the ODE $\hat{\zeta}_t$ evolves independently of the SDE part, while the coefficients of the SDE depend on the evolution of the ODE. We will informally refer to such a

[5] By definition of integral manifold, on each one of these manifolds the rank of the distribution $\hat{\Delta}_0$ is constant and it is equal to the dimension of the manifold itself.

[6] To make a link with the more precise notation that we will use in section 3.1.2 later on, we are denoting here by $\hat{\zeta}_t$ the components (ζ_t, a_t) in (3.8)-(3.9), i.e. $\hat{\zeta}_t = (\zeta_t, a_t)$.

representation as being of the form "ODE+SDE". In general, this representation is only local. This change of coordinates has been known for a long time in differential geometry; here we are simply expressing it in a way which is more congenial to our setting and purposes and we apply it to SDEs. While the change of coordinates itself is not new, for example it has been used by Bismut in [36, Section 5], to study the density of SDEs that satisfy the Hörmander Condition (HC), but to the best of our knowledge it has never been used to study the long-time behaviour of SDEs. This local representation is both an important technical tool throughout this thesis and a fundamental element in understanding the evolution of the dynamics. There is representation has similarities to the Doss-Sussmann method in SDE theory, see [37, Section V.28], which under suitable conditions, expresses the solution of an SDE as $X_t = F(Y_t, B_t)$ where Y_t is the solution of an ODE whose coefficients may be random and F is a deterministic function. It is important for our purposes that the ODE component does not depend on the SDE component so is truly deterministic. Referring to the PDE (1.7), we also note here that the change of coordinates gives a geometric interpretation of the (potential) loss of smoothness in the direction V_0 and of the reason why smoothness is instead maintained in. the direction $\mathcal{V}$, see Note 3.1.13 on this point.

In view of the discussed change of coordinates, it makes sense to start by studying UFG dynamics for which the representation (1.13)-(1.14) is global. For this reason in Section 5.1 we consider systems which are (globally) of the form (1.13)-(1.14), where the ODE is assumed to be one-dimensional and the SDE satisfies a form of Hörmander condition. More precisely, the dynamics studied in Section 5.1 are non-autonomous hypoelliptic SDEs; because the topic is somewhat of independent interest, this section has been written in such a way that it can be read independently of the rest of the thesis. Non-autonomous SDEs and their associated two-parameter semigroup have been studied in [38], where a detailed analysis of the law of the process is carried out, in [39] where the associated semigroup is examined, and in [40, 41], where the authors introduce interesting techniques to deal with the analysis of invariant measures and long-time behaviour of time-inhomogeneous processes. The work [38] assumes that the non-autonomous SDE is hypoelliptic, while in [40] a uniform ellipticity assumption is enforced. From a technical point of view, the results of Section 5.1 extend the approach of [40, Section 6.1] to the hypoelliptic setting. However the main difference between our results and the results in [40] is that here we highlight the fact that the process may admit several invariant measures and we characterize the basin of attraction of each of them. In this setting convergence to equilibrium is driven by the ODE component. We will indeed show that the basin of attraction of each invariant measure can be completely described by just looking at the behaviour of the solution of the ODE. Because the ODE is assumed to be one-dimensional and autonomous, it can only behave monotonically, so the analysis of the ODE and of the full problem is relatively intuitive in this setting (see Section 5.1 for details).

In Section 5.2 we consider the general case of UFG processes for which the representation (1.13)-(1.14) is only local. While this case is substantially richer than the previous one, the fact that, locally, we can always represent the SDE (1.1) as a system of the form "ODE+SDE", still means that there is some deterministic behaviour which is intrinsic to the UFG dynamics. It turns out that one is still able to single out the deterministic behaviour. Recalling the definition of the vector $V_0^{(\perp)}$, formula (1.12), we will show that the (N-dimensional) ODE

$$d\zeta_t = V_0^{(\perp)}(\zeta_t)dt$$

plays, in this more general context, the same driving role that the ODE (1.14) had in the context of Section 5.1. Motivated by the above discussion, we introduce the process

$$\mathcal{Z}_t := e^{-tV_0^{(\perp)}}(X_t).$$

This process is non-autonomous and, as we will explain, it can be interpreted geometrically as being a projection of the process X_t on an appropriate integral manifold of the distribution $\hat{\Delta}$. We apply the techniques of Section 5.1 to the study of such a non-autonomous process, producing results on the long-time behaviour of $\mathcal{Z}_t$. We then relate the asymptotic behaviour of $\mathcal{Z}_t$ to the asymptotic behaviour of X_t. Notice that the procedure that we have just described is somewhat the reverse of the one that is traditionally used (and it is, to the best of our knowledge, new): given a non-autonomous system, the established methodology consists of increasing the dimension of the state space by adding time as an auxiliary variable, thereby reducing the given non-autonomous system to a (larger) autonomous one. Here we do the converse: by projecting the process on an appropriate manifold, we reduce to a (lower-dimensional) non-autonomous one, $\mathcal{Z}_t$, with the advantage that now the techniques of Section 5.1 can be adapted to prove statements on $\mathcal{Z}_t$. Once the latter process has been understood, we deduce results about the autonomous process X_t from those shown for $\mathcal{Z}_t$.

From a probabilistic point of view it is clear that, in the absence of the Hörmander condition, we cannot expect the process X_t to have a density with respect to the Lebesgue measure on $\mathbb{R}^N$. This is made explicit by the local representation (1.13)-(1.14), which also clarifies that it is the ODE component to be responsible for the lack of smoothness. Notice also that, in the coordinates (1.13)-(1.14), the vector $V_0^{(\perp)}$ is given by $V_0^{(\perp)} = (0, \dots, 0, \hat{W}_0)$, i.e. it is precisely the vector driving the ODE behaviour (we have elaborated on this fact in Note 3.1.13). However in Chapter 6 we show that the law of the SDE (1.1) still has a density on an appropriate time-dependent submanifold, which can be explicitly described. In order to do so, we correct and then extend the results of [42].

Chapter 4 is devoted to the study of derivative estimates for Markov semigroups under very relaxed conditions on the coefficients of the SDE. To simplify the discussion we assume for the moment, that $N = d = 1$ in (1.1), and we investigate conditions to ensure that the derivative of the semigroup $\mathcal{P}_t$ in the direction V_1, i.e. $V_1\mathcal{P}_t f$, decays exponentially fast to zero as t tends to ∞. In this thesis we will prove such estimates in the one dimensional situation for convenience, in [2] we generalise to a multidimensional situation, we aim to cover the full problem in [43].

The study of derivative estimates for Markov semigroups has a long history and it has been tackled by a multitude of more or less general approaches [25, 44–47], to mention just a few. As is well-known, without any quantitative condition on the vector fields appearing in (1.1) (i.e. if only ellipticity/hypoellipticity or other regularity assumptions are made), only the following smoothing-type estimates hold

$$|V_1\mathcal{P}_t f(x)| \leq u(x)\frac{1}{t^\gamma}, \quad \forall t \in (0, 1),$$

where $\gamma > 0$ is an appropriate exponent depending on which field V we are differentiating along, see [25, 44, 47–50], and most of the literature is devoted to estimates of the above type. In [32] the authors introduced a sufficient condition which they named the *Obtuse Angle Condition* (OAC), in order for the following estimate to hold: for every $r > 0$ and $t_0 > 0$, we may find a constant $c_{t_0,r} > 0$ such that for any $f \in C_b(\mathbb{R})$ and $t \geq t_0$ we have

$$\sup_{\mathbb{B}(0,r)} |V_1(\mathcal{P}_t f)(x)| \leq c_{r,t_0} \|f\|_\infty e^{-\overline{\lambda}(t-t_0)}, \tag{1.15}$$

where $\mathbb{B}(0,r)$ is the Euclidean ball with $\mathbb{R}^N$ of radius r and centre 0. In this simplified setting, the OAC can be expressed as follows: there exists a constant $\lambda > 0$ such that

$$[V_1, V_0](x) V_1(x) \leq -\lambda |V_1(x)|^2, \quad \forall x \in \mathbb{R}. \tag{1.16}$$

This is a coercivity-type condition (inspired by dilation structures in Carnot groups, see [51]); in the above such coercivity is required to hold uniformly in space in the sense that $\lambda > 0$ is a constant independent of x. In contrast, in Chapter 4 we discuss the case in which λ is allowed to be a measurable function of x. That is, we consider the following condition: there exists a function $\lambda : \mathbb{R} \to \mathbb{R}$ such that

$$([V_1, V_0](x))\, V_1(x) \leq -\lambda(x) |V_1(x)|^2, \quad \forall x \in \mathbb{R}\,, \tag{1.17}$$

and we refer to (1.17) as the *Local Obtuse Angle Condition* (LOAC). In Section 4.1 we give a simple example to further explain why we name (1.17) the *local* OAC, see comments after equation (4.5). Under no further assumptions on the function $\lambda = \lambda(x)$ (neither on the regularity nor on the sign of such a function) we show that the following holds

$$|V_1 \mathcal{P}_t f(x)| \leq c\, \mathbb{E}\left[\exp\left(-2 \int_0^t \lambda(X_s) ds\right)\right]^{\frac{1}{2}}, \tag{1.18}$$

for some constant $c > 0$. In order to obtain estimates of the form (1.15) and (1.18) under the local condition (1.17), we need to gain more detailed control over the paths of the diffusion X_t; for this reason we initiate in this thesis a pathwise version of the Bakry-Emery approach [25] to the study of derivative estimates for Markov semigroups. This is the content of Section 4.1 (and in Note 4.1.9 we explain why classical arguments do not carry through in this setting). Clearly, if $\lambda(x) \geq \lambda_0 > 0$ for some constant λ_0 then (1.15) follows from (1.18). If $\lambda(x)$ is not uniformly bounded below, one can still obtain (1.15) from (1.18). This is what we show in Section 4.2. Roughly speaking, in Section 4.2 we show that if there exists a set F such that $\lambda(x) \geq \lambda_0 > 0$ for every $x \in F$ and the processs spends enough time in such a set, then one can still obtain (1.15). Such an approach allows us to treat SDEs with bounded coefficients, which do not satisfy Lyapunov conditions and for which one would not necessarily expect (1.15) see Section 4.1. In order to obtain such results we make use of Large Deviation principles; in particular, we use (and weaken) some estimates on functionals of the occupation measure which have been obtained by Donsker and Varadhan in [52–55]. This provides a link between the study of derivative estimates for Markov Semigroups and Large Deviations theory and allows one to give an explicit characterization of the dependence on r on the right hand side of (1.15).

1.3 Organisation of the thesis

In Section 2.1 we introduce the standing notation for the remainder of the manuscript. To make the thesis self-contained, in Chapter 2 we gather background definitions and notions. In particular Section 2.2 contains details about the UFG condition, while Section 2.3 covers basic definitions and standard results in differential geometry and (stochastic) control theory. In Section 3.1 we exploit the existing theory of distributions of non-constant rank to produce both global and local results about the SDE (1.1), under the UFG condition. In Section 3.1.1 we cover the *global* behaviour of the SDE, in Section 3.1.2 we study *local* properties. In Section 3.2 we introduce several results for UFG-diffusions. These results are quite general, in the sense that most of them are valid under just the UFG condition. Section 5.1 can be read independently of the rest of the manuscript: in this section we describe the long-time behaviour of hypoelliptic SDEs of non-autonomous type. The class of SDEs considered in Section 5.1 is one for which the representation of the form "ODE+SDE" is global. This is the first section where we address the problem of studying the basin of attraction of different invariant measures. In Section 5.2 we instead study the long time behaviour of (1.1) in the general UFG case (in which the change of coordinates is only local). Chapter 6 is devoted to the study of the density of the process, via Malliavin calculus. Section 4.1 presents the pathwise approach developed to obtain estimates of the type (1.18) from the non-uniform coercivity condition (1.17). This pathwise approach is, to the best of our knowledge, new; to present the main ideas without cumbersome notations and with proofs of contained length, all the results of Section 4.1 are presented in one dimension, i.e. in the case in which the SDE (1.1) lives in $\mathbb{R}$. A generalisation of these results to $\mathbb{R}^N$ with additional assumptions on the commutator structure can be found in [2], more complete extensions to $\mathbb{R}^N$ are lengthy and significantly more technical, and they will be the object of future work, [43]. We explain in Note 4.1.9 the reason why, under such a non-uniform condition, classical approaches can no longer be used. For consistency with Section 4.1, also the results of Section 4.2 are presented in one dimension, however the results of this section are in reality completely dimension-independent. Section 4.3 contains several examples and counterexamples to illustrate cases where the results developed in this thesis apply.

Chapter 2

Preliminaries and Assumptions

2.1 Notation

We will be interested in N-dimensional SDEs, of the form (1.1). The letter N will only be used to refer to the dimension of the state space.

If x is a point in $\mathbb{R}^N$, we denote the j-th coordinate of x by x^j, i.e. $x = (x^1, \dots, x^N)$ (this is coherent with (1.3)). We will often use a local change of coordinates, presented in Section 3.1.2. The change of coordinates will be given by a local diffeomorphism $\Phi : \mathbb{R}^N \to \mathbb{R}^N$ and the new coordinates will typically be denoted by $\mathbf{z}$, i.e. $\mathbf{z} = \Phi(x)$. In the new coordinate system it will be of particular importance to distinguish the role of the first n coordinates of $\mathbf{z}$ from the others (n being an appropriate integer, $n < N$). In particular, if $N - n > 1$, we will use the following notation

$$\mathbf{z} = (\underbrace{z^1, \dots, z^n}_{z}, \underbrace{z^{n+1}}_{\zeta}, \underbrace{z^{n+2}, \dots, z^N}_{a}) = (z, \zeta, a), \tag{2.1}$$

where $(z^1, \dots, z^n) = z \in \mathbb{R}^n$, $z^{n+1} = \zeta \in \mathbb{R}$ and $(z^{n+2}, \dots, z^N) = a \in \mathbb{R}^{N-(n+1)}$. The last block of coordinates plays a role which is different from the role of the first two blocks, as it will be explained (the coordinates in the last block should be more intended as parameters). If $N = n + 1$ then simply $\mathbf{z} = (z, \zeta)$.

A similar reasoning holds for the vector fields appearing in (1.1): for any $j \in \{0, \dots, d\}$, $V_j = (V_j^1, \dots, V_j^N)$ and $\tilde{V}_j$ will denote the vector V_j, expressed in the new coordinate system $\mathbf{z} = \Phi(x)$. We will show that in the new coordinate system, one has

$$\tilde{V}_j(\mathbf{z}) = (U_j(\mathbf{z}), 0, \dots, 0) \qquad j = 1, \dots, d \tag{2.2}$$

$$\tilde{V}_0 = (U_0(\mathbf{z}), W_0(\zeta, a), 0, \dots, 0)\,, \tag{2.3}$$

where $U_j : \mathbb{R}^N \to \mathbb{R}^n$ while W_0 is a real-valued function which depends only on the last two blocks of coordinates of z, i.e. $W_0 : \mathbb{R}^{N-n} \to \mathbb{R}$.

Accordingly, if $\mathbb{R}^N \ni X_t$ is the solution at time t of the SDE (1.1), then X_t^j denotes the j-th component of X_t. We will sometimes want to stress the dependence of the solution X_t on the initial datum; when this is the case, we will write $X_t^{(x)}$ if $X_0 = x$. Finally, given a probability measure μ and a function f which is integrable with respect to μ, we shall define $\mu(f)$ by

$$\mu(f) := \int f d\mu.$$

We shall use the following function spaces throughout the thesis. For any $N \geq 1$ and closed set $E \subseteq \mathbb{R}^N$;

- We denote by $C_b(E)$ the space of all functions $f : E \to \mathbb{R}$ which are continuous and bounded; this space will be endowed with the supremum norm.
- We denote by $C_c^\infty(\mathbb{R}^N)$ the set of all functions $f : \mathbb{R}^N \to \mathbb{R}$ which are C^∞ and with compact support.

Given a differentiable function $f : \mathbb{R}^N \to \mathbb{R}^N$ we denote by $\mathcal{J}_x f$ the Jacobian matrix of f, that is $(\mathcal{J}_x f)^{ij}(x) = \partial_{x^j} f^i(x)$.

2.2 The UFG condition

Fix $d \in \mathbb{N}$ and let $\mathcal{A}$ be the set of all k-tuples, of any size $k \geq 1$, of integers of the following form

$$\mathcal{A} := \{\alpha = (\alpha^1, \dots, \alpha^k), k \in \mathbb{N} : \alpha^j \in \{0, 1, \dots, d\} \text{ for all } j \geq 1\} \setminus \{(0)\}\,.$$

We emphasise that all k-tuples of any length $k \geq 1$ are allowed in $\mathcal{A}$, except the trivial one, $\alpha = (0)$ (however singletons $\alpha = (j)$ belong to $\mathcal{A}$ if $j \in \{1, \ldots, d\}$). We endow $\mathcal{A}$ with the product operation

$$\alpha * \beta := (\alpha^1, \ldots, \alpha^h, \beta^1, \ldots, \beta^\ell),$$

for any $\alpha = (\alpha^1, \ldots, \alpha^h)$ and $\beta = (\beta^1, \ldots, \beta^\ell)$ in $\mathcal{A}$. If $\alpha \in \mathcal{A}$, we define the *length* of α, denoted by $\|\alpha\|$, to be the integer

$$\|\alpha\| := h + \text{card}\{i : \alpha_i = 0\}, \qquad \text{if } \alpha = (\alpha^1, \ldots, \alpha^h)\,.$$

For any $m \in \mathbb{N}, m \geq 1$, we then introduce the sets

$$\mathcal{A}_m = \{\alpha \in \mathcal{A} : \|\alpha\| \leq m\}\,.$$

Given a vector field (or, equivalently, a first order differential operator) $V = (V^1(x), V^2(x), ..., V^N(x))$ on $\mathbb{R}^N$, we refer to the functions $\{V^j(x)\}_{1 \leq j \leq N}$ as to the *components* or *coefficients* of the vector field. We say that a vector field is smooth or that it is C^∞ if all the components $V^j(x)$, $j = 1, \ldots, N$, are C^∞ functions. Given two differential operators V and W, the commutator between V and W is defined as

$$[V, W] := VW - WV\,.$$

Let now $\{V_i : i = 0, \ldots, d\}$ be a collection of vector fields on $\mathbb{R}^N$ and let us define the following "hierarchy" of operators:

$$V_{[i]} := V_i \qquad i = 0, 1, \ldots, d$$

$$V_{[\alpha * i]} := [V_{[\alpha]}, V_{[i]}], \qquad \alpha \in \mathcal{A}, i = 0, 1, \ldots, d\,.$$

This hierarchy is completely analogous to the one constructed in the Introduction, here we just need a more detailed notation. Note that if $\|\alpha\| = h$ then $\|\alpha * i\| = h+1$ if $i \in \{1, \ldots, d\}$ and $\|\alpha * i\| = h+2$ if $i = 0$. If $\alpha \in \mathcal{A}$ is a multi-number of length h, with abuse of nomenclature we will say that $V_{[\alpha]}$ is a differential operator of length

h. We can then define the space $\mathcal{R}_m$ to be the space containing all the operators of the above hierarchy, up to and including the operators of length m (but excluding V_0):

$$\mathcal{R}_m := \left\{V_{[\alpha]}, \alpha \in \mathcal{A}_m\right\}. \tag{2.4}$$

Let $C_V^\infty(\mathbb{R}^N)$ denote the set of bounded smooth functions, $\varphi = \varphi(x) : \mathbb{R}^N \to \mathbb{R}$, such that

$$\sup_{x\in\mathbb{R}^N} \left|V_{[\gamma_1]} \dots V_{[\gamma_k]}\varphi\right| < \infty \tag{2.5}$$

for all k and all $\gamma_1, \dots, \gamma_k \in \mathcal{A}_m$. With this notation in place we can now introduce the definition that will be central in this thesis.

Definition 2.2.1 (UFG Condition)**.** Let $\{V_i : i = 0, \dots, d\}$ be a collection of smooth vector fields on $\mathbb{R}^N$ and assume that the coefficients of such vector fields have bounded partial derivatives (of any order). We say that the vector fields $\{V_i : i = 0, \dots, d\}$ satisfy the UFG condition if there exists $m \in \mathbb{N}$ such that for any $\alpha \in \mathcal{A}$ of the form

$$\alpha = \alpha' * i, \qquad \alpha' \in \mathcal{A}_m,\, i \in \{0, \dots, d\},$$

one can find bounded smooth functions $\varphi_{\alpha,\beta} = \varphi_{\alpha,\beta}(x) \in C_V^\infty(\mathbb{R}^N)$ such that

$$V_{[\alpha]}(x) = \sum_{\beta\in\mathcal{A}_m} \varphi_{\alpha,\beta}(x)V_{[\beta]}(x)\,. \tag{\textbf{UFG}}$$

Again we emphasize that the set of vector fields appearing in the linear combination on the right hand side of the above identity does not include V_0. It may be useful to compare the UFG condition with the Hörmander condition (HC), the Uniform Parabolic Hörmander condition (UPHC) and the Parabolic Hörmander condition (PHC), which we recall. The HC is satisfied if

$$\text{Lie}\{V_0(x), \dots, V_d(x)\} = \mathbb{R}^N \quad \text{for every } x \in \mathbb{R}^N\,. \tag{\textbf{HC}}$$

The PHC has been recalled in the introduction, see (**PHC**). We notice in passing that while the space $\mathcal{R}_m$ is in general different from the space $\mathfrak{L}_m$, it is the case that

$\cup_{j\geq 1}\mathcal{R}_j = \cup_{j\geq 1}\mathfrak{L}_j$. The UPHC (see [56]) is instead satisfied if

$$\exists\, \ell \geq 1 \text{ and } \kappa > 0 : \sum_{\alpha\in\mathcal{A}_\ell} \left|V_{[\alpha]}(x)\cdot y\right|^2 \geq \kappa\,|y|^2 \quad \text{for every } x, y \in \mathbb{R}^N . \qquad \textbf{(UPHC)}$$

In the above each term of the sum is the scalar product between the vector $V_{[\alpha]}(x)$ and the vector $y \in \mathbb{R}^N$. Notice that the UPHC is the strongest of all these conditions, in the sense that

$$(\textbf{UPHC}) \Rightarrow (\textbf{PHC}) \Rightarrow (\textbf{HC}) \tag{2.6}$$

$$(\textbf{UPHC}) \Rightarrow (\textbf{UFG})\,. \tag{2.7}$$

However neither the HC nor the PHC imply the UFG condition. We also note that while the various Hörmander conditions are imposed on an appropriate Lie Algebra, the UFG condition is rather a condition on the set of vectors $\{V_{[\alpha]}, \alpha \in \mathcal{A}_m\}$, seen as a module over the ring C_V^∞. The implications of (2.6) and (2.7) will be shown to hold true in Appendix A.1.2. For completeness we recall that the ellipticity condition is satisfied if

$$\exists \kappa > 0 : \sum_{i=1}^{d} |V_i(x)\cdot y|^2 \geq \kappa\,|y|^2 \quad \text{for every } x, y \in \mathbb{R}^N .$$

Note 2.2.2 (On nomenclature). Throughout the thesis we shall say the vector fields $V_0, V_1, \dots, V_d$ satisfy the parabolic Hörmander condition (UFG condition respectively), the operator $\mathcal{L}$ satisfies the parabolic Hörmander condition (UFG condition respectively) and the process X_t is a Hörmander type SDE (UFG type SDE respectively) interchangeably.

Example 2.2.3. Recall the one-dimensional geometric Brownian motion (1.5), that is the case when $N = d = 1$ and $V_0 = V_1 = x\partial_x$. As was shown in Section 1.1 for this example (**PHC**) is not satisfied. However, these vector fields satisfy the UFG condition with $m = 1$ as $[V_1, V_0] = 0$. □

Example 2.2.4. Consider the following first order differential operators on $\mathbb{R}^2$

$$V_0 = \sin x\,\partial_y \qquad V_1 = \sin x\,\partial_x\,.$$

Then $\{V_0, V_1\}$ do not satisfy the Hörmander condition (e.g. there is always a degeneracy at $x = 0$) but they do satisfy the UFG condition with $m = 4$. If the role of the fields is exchanged, i.e. if we set

$$V_0 = \sin x\, \partial_x, \qquad V_1 = \sin x\, \partial_y$$

then $\{V_0, V_1\}$ still satisfy the UFG condition, this time with $m = 1$ (indeed, $[V_0, V_1] = \cos(x)V_1$). □

Note 2.2.5. Because the functions $\varphi_{\alpha,\beta}$ appearing in (**UFG**) belong to $C_V^\infty(\mathbb{R}^N)$, if the UFG condition holds for some $m \in \mathbb{N}$ then it also holds for any $\ell \geq m, \ell \in \mathbb{N}$. In other words, if the UFG condition holds for some m in $\mathbb{N}$ then for any $V_{[\gamma]}$ with $\|\gamma\| > m$ one has

$$V_{[\gamma]}(x) = \sum_{\beta \in \mathcal{A}_m} \varphi_{\gamma,\beta}(x) V_{[\beta]}(x)\,,$$

for some functions $\varphi_{\gamma,\beta} \in C_V^\infty(\mathbb{R}^N)$. See Lemma A.1.7 for details. For this reason it is appropriate to remark that in the remainder of the thesis, when we assume that "the UFG condition is satisfied for some m", we mean the smallest such m. □

We will consider diffusion semigroups $\{\mathcal{P}_t\}_{t\geq 0}$ of the form (1.2); that is, we consider Markov semigroups associated with the stochastic dynamics (1.1). In particular, we will be interested in studying the semigroup $\mathcal{P}_t$ when the vector fields $\{V_0, V_1, \ldots, V_d\}$ satisfy the UFG condition.

As we have already mentioned, the UFG condition is strictly weaker than the Uniform Parabolic Hörmander condition. However one can still prove that, when such a condition is satisfied by the vector fields $\{V_0, V_1, \ldots, V_d\}$ appearing in the generator (1.4), the semigroup $\mathcal{P}_t$ still enjoys good smoothing properties: if $f(x)$ is continuous then $(\mathcal{P}_t f)(x)$ is differentiable (infinitely many times) in all the directions spanned by the vector fields contained in $\mathcal{R}_m$ (we recall that the set $\mathcal{R}_m$ is defined in (2.4)). See Appendix A.1.4 for more details.

When the semigroup $\mathcal{P}_t$ is elliptic or hypoelliptic, several works have dealt with the study of the long and short time behaviour of the derivatives of the semigroup, for a review see [8, 25]. To the best of our knowledge, the only work addressing the study of the long-time behaviour of the derivatives of UFG semigroups is [32]. In [32] the authors identify a sufficient condition for exponential decay of the derivatives of the solution of (1.7). To be more precise, they proved the following: suppose the vector fields $\{V_0, V_1, \ldots, V_d\}$ satisfy the UFG condition and assume there exists $\lambda_0 > 0$ such that for all f sufficiently smooth and for every $\alpha \in \mathcal{A}_m$ we have

$$(V_{[\alpha*0]}f(x))(V_{[\alpha]}f(x)) \leq -\lambda_0 \left|V_{[\alpha]}f(x)\right|^2, \quad \text{for every } x \in \mathbb{R}^N. \tag{2.8}$$

If λ_0 is sufficiently large then, for every $r > 0$ and $t_0 > 0$, we may find a constant $c_{t_0,r} > 0$ such that for any $f \in C_b(\mathbb{R})$, $t \geq t_0$ and $\alpha \in \mathcal{A}_m$ we have

$$\sup_{\mathbb{B}(0,r)} \left|V_{[\alpha]}(\mathcal{P}_t f)(x)\right| \leq c_{r,t_0}\|f\|_\infty e^{-\lambda(t-t_0)}, \tag{2.9}$$

for some $\lambda > 0$. In the above $\mathbb{B}(0,r)$ is the centered ball (of $\mathbb{R}^N$) of radius r. Condition (2.8) was named the *Obtuse Angle Condition* (OAC) in [32]. In Chapter 4 we give weaker conditions under which (2.9) holds in a one-dimensional setting. Here we will need a second order version of such a result, as well.

Lemma 2.2.6. *Let $\mathcal{P}_t$ be the semigroup associated with the SDE* (1.1) *and assume that the vector fields $V_0, \ldots, V_d$ satisfy the UFG condition. Suppose moreover that the following holds: there exists $\lambda_0 > 0$ such that*

$$(V_{[\alpha]}V_{[\beta]}f)(x)\,([V_{[\alpha]}V_{[\beta]}, V_0]f)(x) \leq -\lambda_0 \left|(V_{[\alpha]}V_{[\beta]}f)(x)\right|^2, \tag{2.10}$$

for every $x \in \mathbb{R}^N$ and for all $\alpha, \beta \in \mathcal{A}_m$ such that $\alpha \neq \beta$ and $\alpha, \beta \notin \{1, \ldots, d\}$. If $\lambda_0 > 0$ is large enough then, for any $t_0 \in (0,1)$ and any $r > 0$ there exists a constant $c_{t_0,r} > 0$ such that, for some $\lambda > 0$, one has

$$\sup_{x\in\mathbb{B}(0,r)} \left|V_{[\beta]}V_{[\alpha]}(\mathcal{P}_t f)(x)\right|^2 \leq c_{t_0,r}\, e^{-\lambda(t-t_0)}\|f\|_\infty, \tag{2.11}$$

for all $\alpha, \beta \in \mathcal{A}_m$, all $t > t_0$ and for every f continuous and bounded.

Example 2.2.7 (UFG condition and Obtuse Angle Condition for linear SDEs)**.** Consider SDEs in $\mathbb{R}^N$ of the form

$$dX_t = (AX_t + D)dt + \sqrt{2}\sum_{i=1}^{d} C_i dB_t^i\,, \tag{2.12}$$

where A is a constant $N \times N$ matrix, $B_t^1, \ldots, B_t^d$ are one-dimensional standard Brownian motions and $D, C_1, \ldots, C_d \in \mathbb{R}^N$ are constant vectors. In this case $V_0(x) = Ax + D, V_i(x) = C_i$, and

$$V_{[i*0]} = [V_i, V_0] = AC_i, \quad i \in \{1, \ldots, d\}.$$

Because $[V_i, V_j] = 0$ for every $i, j \in \{1, \ldots, d\}$, the only relevant commutators are those of the form $V_{[(i,0,\ldots,0)]}$, i.e. repeated commutators with V_0. For simplicity, let $\alpha_{i,k}$ be the $(k+1)$-tuple such that $\alpha_{i,k}^1 = i$ and $\alpha_{i,k}^j = 0$ for $j > 1$; then

$$V_{[(i,\underbrace{0,\ldots,0}_{k \text{ times}})]} = A^k C_i.$$

It is now easy to show that, *irrespective of the choice of $A, D, C_1, \ldots, C_d$ as above, the UFG condition is always satisfied by SDEs of the form* (2.12). Indeed, by the Cayley Hamilton Theorem there is a polynomial p of degree at most $N - 1$ such that $A^N = p(A)$; so we can write any $V_{[\alpha_{i,k}]}$ as a linear combination of the vectors $V_{[\alpha_{i,\ell}]}$ with $\ell \leq N$. For comparison we recall that (2.12) is hypoelliptic if and only if the Kalman rank condition is satisfied, namely if

$$\operatorname{rank}[Q, AQ, A^2Q, \ldots, A^{N-1}Q] = N,$$

where Q is the overall diffusion matrix of (2.12), see e.g. [8]. As for the OAC (2.8), this is is satisfied if and only if there exists some $\lambda_0 > 0$ such that for all f sufficiently smooth we have

$$(\nabla f)^T A^{k+1} C_i C_i^T (A^k)^T \nabla f \leq -\lambda (\nabla f)^T A^k C_i C_i^T (A^k)^T (\nabla f), \tag{2.13}$$

for all $i \in \{1, \ldots, d\}$ and $k \in \{0, 1, \ldots, N - 1\}$. Equivalently this holds if and only if

$$(A + \lambda I)B \leq 0,$$

where $B = A^k C_i C_i^T (A^k)^T$ for all $i \in \{1, \ldots, d\}$ and $k \in \{0, 1, \ldots, N - 1\}$. $\square$

2.3 Background Geometry

In this section we cover some basic notions from differential geometry and geometric control theory on which the rest of the thesis relies. Further details can be found in the excellent references [20, 57, 58]. For the reader who is already familiar with this material, we point out that, among the results recalled in this section, Theorem 2.3.6 is possibly the one which will play the most important role in the remainder of the thesis.

Given a vector field $V(x)$ on $\mathbb{R}^N$, we denote by $e^{tV}(x)$ the integral curve of V starting at $t = 0$ from x, i.e. the curve $\gamma : \mathbb{R} \to \mathbb{R}^N$ such that $\gamma(0) = x$ and $\dot{\gamma}(t) = V(\gamma(t))$ for all $t \in \mathbb{R}$ such that the curve is defined. In general, integral curves exist only locally. In this thesis we consider only smooth, globally defined and globally Lipschitz vector fields (see Hypothesis 2.4.1), so integral curves actually exist for every $t \in \mathbb{R}$. As already mentioned, a *distribution* Δ on $\mathbb{R}^N$ is a map that, to each point $x \in \mathbb{R}^N$, associates a linear subspace of the tangent space $T_x\mathbb{R}^N$. Given a set $\mathcal{D}$ of smooth vector fields on $\mathbb{R}^N$, the distribution generated by $\mathcal{D}$, denoted by $\Delta_{\mathcal{D}}$, is the map $x \to \mathrm{span}\{V(x) : V \in \mathcal{D}\}$. Distributions generated by a set of smooth vector fields are usually referred to as *smooth distributions.* When we write $\Delta_{\mathcal{D}}$ instead of just Δ it is understood that we are considering smooth distributions rather than general distributions. As customary, we say that the vector field V on $\mathbb{R}^N$ belongs to the distribution Δ if $V(x) \in \Delta(x)$ for all $x \in \mathbb{R}^N$. The *rank* of Δ at x is the dimension of the vector space $\Delta(x)$. A *piecewise integral curve*, γ, of vector fields in the set $\mathcal{D}$ is a curve of the form

$$\gamma(t_1, \ldots, t_h) = e^{t_1 X_1} e^{t_2 X_2} \cdots e^{t_h X_h} x \qquad h \in \mathbb{N}, t_j \in \mathbb{R}, x \in \mathbb{R}^N,$$

where $X_1, \ldots, X_h \in \mathcal{D}$ (and they are not necessarily all distinct). A submanifold $M \subseteq \mathbb{R}^N$ is an *integral manifold* of Δ if $T_x M = \Delta(x)$ for every $x \in M$. A *maximal integral manifold* (MIM) of Δ, $\mathcal{M}$, is a connected integral manifold of Δ which is maximal in the sense that every other connected integral manifold of Δ that contains $\mathcal{M}$ coincides with $\mathcal{M}$. Therefore, two MIMs either coincide or they are disjoint.

Definition 2.3.1. Let Δ be a distribution on $\mathbb{R}^N$.

- Δ is *involutive* if
$$X, Y \in \Delta \quad \Longrightarrow \quad [X, Y] \in \Delta\,.$$

- Δ is invariant under the vector field V if the Jacobian matrix $\mathcal{J}_x(e^{tV}x)$ maps $\Delta(x)$ into $\Delta(e^{tV}x)$ for all x and for all t.[7]

- Suppose Δ is generated by the collection of vector fields $\mathcal{D} = \{X_1, \ldots, X_k\}$, i.e. $\Delta = \Delta_{\mathcal{D}}$. Then two points $x, y \in \mathbb{R}^N$ belong to the same *orbit* of $\Delta_{\mathcal{D}}$ if there exists a curve $\gamma : [a, b] \to \mathbb{R}^N$ such that $\gamma(a) = x$, $\gamma(b) = y$ and γ is a piecewise integral curve of vectors in $\mathcal{D}$.

In general, the integral manifolds of a given distribution are "smaller" than the orbits; we refer the reader to [20] for a detailed explanation of this matter, see in particular [20, Eqn. (3.1)]. Here we just illustrate this fact with a simple but important classical example.

Example 2.3.2. In $\mathbb{R}^2$, consider the vector fields $X = \partial_x$ and $Y = \psi(x)\partial_y$ where $\psi(x)$ is a smooth function vanishing on the half-plane $x \leq 0$. The orbit of the distribution generated by X and Y, $\Delta_{X,Y}$, is the whole $\mathbb{R}^2$. That is, given any two points in $\mathbb{R}^2$ there is a piecewise integral curve of $\{X, Y\}$ which joins the two points. However the integral manifolds through points (x, y) with $x \leq 0$ are one dimensional. Notice that the distribution in this example is involutive but it satisfies neither the Hörmander condition nor the UFG condition. More precisely, in the sense that whether we take $X = V_0$ and $Y = V_1$ or vicecersa, either ways the UFG condition is not satisfied (more precisely, in the language of Definition 2.3.4 below, the set $\{X, Y\}$ is neither locally nor globally of finite type). The fact that $\{X, Y\}$ don't satisfy the UFG condition can be either seen as a consequence of Theorem 2.3.6 below (if it did, the orbits would have to coincide with the integral manifolds) or it can be shown with direct calculations (the problem arising on the line $x = 0$). For the reader's convenience this calculation is contained in the Appendix, see Lemma A.1.13. $\square$

We say that a distribution Δ on $\mathbb{R}^N$ satisfies the *(maximal) integral manifolds property* if through every point of $\mathbb{R}^N$ there passes a (maximal) integral manifold

[7] A useful criterion to check whether a distribution is invariant under a vector field will be given in Note 2.3.5.

of Δ. The following fundamental result, due to Sussmann (see [20, Theorem 4.2]), completely characterizes the distributions enjoying the maximal integral manifolds property.

Theorem 2.3.3 (Sussmann's Orbit Theorem)**.** *If $\Delta = \Delta_{\mathcal{D}}$ is a smooth distribution on $\mathbb{R}^N$, then the following statements are equivalent*

(a) *$\Delta_{\mathcal{D}}$ satisfies the maximal integral manifolds property;*

(b) *$\Delta_{\mathcal{D}}$ satisfies the integral manifolds property;*

(c) *the orbits of $\Delta_{\mathcal{D}}$ coincide with the integral manifolds of the distribution and the rank of $\Delta_{\mathcal{D}}$ at each point $x \in \mathbb{R}^N$ is equal to the dimension of the integral manifold of $\Delta_{\mathcal{D}}$ through x;*

(d) *$\Delta_{\mathcal{D}}$ coincides with the smallest distribution which contains the Lie algebra generated by $\mathcal{D}$, Lie$\{\mathcal{D}\}$, and is invariant under the vectors in $\mathcal{D}$.*

In view of the equivalence of (a) and (b) above, when either property hold we just say that the smooth distribution is *integrable.* It is clear that in this case every integral manifold is a maximal integral manifold. Some standard facts about integrable distributions which are useful to bear in mind and that follow (easily) from what we have said so far: if $\Delta_{\mathcal{D}}$ is integrable, then

i) $\Delta_{\mathcal{D}}$ is involutive;

ii) the state space $\mathbb{R}^N$ is partitioned into orbits of $\Delta_{\mathcal{D}}$;

iii) the rank of the distribution is constant along the orbits (of $\Delta_{\mathcal{D}}$, which coincide with the integral manifolds of such a distribution).

The latter fact is a consequence of the fact that $\Delta_{\mathcal{D}}$ is invariant under the vectors in $\mathcal{D}$ together with the observation that the maps e^{tV} are diffeomorphisms for every fixed $t \in \mathbb{R}$ (hence the Jacobian matrix $\mathcal{J}_x(e^{tV}x)$, which maps the tangent space at x into the tangent space at $e^{tV}x$, is always invertible).

Definition 2.3.4 ([20, page 185])**.** Let $\mathcal{D}$ be a set of everywhere defined, smooth vector fields on $\mathbb{R}^N$ and $\Delta_{\mathcal{D}}$ be the associated distribution. The set $\mathcal{D}$ (as well as the distribution $\Delta_{\mathcal{D}}$) is *locally of finite type* or *locally finitely generated* (LFG) if for every $x \in \mathbb{R}^N$ there exist vector fields $X_1, \ldots, X_k$ such that

i) $\operatorname{span}\{X_1(x), \ldots, X_k(x)\} = \Delta_{\mathcal{D}}(x)$

ii) for every $X \in \mathcal{D}$ there exists a neighbourhood U of x and C^∞ functions $\varphi_{i,j}$ defined on U such that

$$[X, X_i] = \sum_{j=1}^{k} \varphi_{i,j}(x) X_j(x) \quad \text{for all } x \in U \text{and every } i \in \{1, \ldots, k\}.$$

We emphasize that if $\Delta_{\mathcal{D}}$ is LFG then the rank of $\Delta_{\mathcal{D}}$ need not be constant.

Note 2.3.5. We recall the following useful criterion (see [58, Lemma 2.1.4]): if a distribution Δ is either of constant rank or locally of finite type, then it is invariant under a vector field V if and only if $[V, \tau] \in \Delta$ whenever $\tau \in \Delta$. □

The next theorem gives a sufficient condition for integrability, which is easy to check in practice.

Theorem 2.3.6 (Hermann, Lobry, Stephan and Sussmann)**.** *If $\mathcal{D}$ is locally of finite type then $\Delta_{\mathcal{D}}$ is integrable; in particular, the integral manifolds of $\Delta_{\mathcal{D}}$ coincide with the orbits of the vector fields of the set $\mathcal{D}$.*

Note 2.3.7 (Comments on Theorem 2.3.6). Seen from a control-theoretical point of view, the above statement gives a global decomposition of the state space $\mathbb{R}^N$ into sets reachable by piecewise integral curves of vector fields in $\mathcal{D}$. To clarify this fact and provide some context, it is useful to compare it with the case where the HC holds. Start by noting that under the HC the Lie algebra generated by the vectors in $\mathcal{D}$ is required to have constant rank (and the rank is assumed to be precisely N at every point). The control-theoretical meaning of the HC is expressed by Chow's Theorem, see [51, 59], (and indeed in control theory the HC is known as *Chow's condition*). Chow's theorem states that if the vectors $\{V_0, \ldots, V_d\}$ satisfy (**HC**) then any two points in $\mathbb{R}^N$ are *accessible* or *reachable* in finite time from each other along integral curves of the vectors in $\mathcal{D}$. That is, given any two points $x, y \in \mathbb{R}^N$, there exists a piecewise integral curve γ of vectors in $\mathcal{D}$, and a time $t > 0$ such that $\gamma(0) = x$ and $\gamma(t) = y$. This is not the case if we simply assume that $\mathcal{D}$ is a LFG set of vector fields. According to the above theorem, if $\mathcal{D}$ is LFG then, for every $x \in \mathbb{R}^N$, the set of states reachable from x in finite time coincides with the maximal integral manifold of $\Delta_{\mathcal{D}}$ through x. Because the rank of the distribution is not constant, and in particular it needs not be N at any point, this implies that, in general, the orbits of $\Delta_{\mathcal{D}}$ will be proper subsets of $\mathbb{R}^N$ (as we have mentioned, they form a partition of $\mathbb{R}^N$). □

We conclude this subsection by recalling the following result, which will be used later on.

Lemma 2.3.8 ([58, Theorem 2.1.9]). *Let $\Delta_{\mathcal{D}}$ be a smooth involutive distribution invariant under a vector field W. Suppose $\Delta_{\mathcal{D}}$ is locally finitely generated. Let x_1, x_2 be two points belonging to the same maximal integral manifold of $\Delta_{\mathcal{D}}$. Then, for all $t \in \mathbb{R}$, the points $e^{tW}x_1$ and $e^{tW}x_2$ belong to the same maximal integral submanifold of $\Delta_{\mathcal{D}}$.*

To clarify the above statement: under the asumptions of the lemma, if x_1, x_2 belong to a given MIM of $\Delta_{\mathcal{D}}$, say $\mathcal{M}$ then $e^{tW}x_1, e^{tW}x_2 \in \tilde{\mathcal{M}}$, where $\tilde{\mathcal{M}}$ denotes another generic MIM of $\Delta_{\mathcal{D}}$. In general $\tilde{\mathcal{M}}$ will be different from $\mathcal{M}$ (unless W belongs to $\Delta_{\mathcal{D}}$). For example see Example 3.1.10.

2.4 Standing Assumptions

Throughout the thesis we will make the following standing assumptions.

Hypothesis 2.4.1. Standing assumptions:

[**SA.1**] All the vector fields we consider in this thesis are smooth, everywhere defined and globally Lipschitz.

[**SA.2**] In this thesis we will consider partitions of $\mathbb{R}^N$ into submanifolds; each one of such submanifolds is generically denoted by $\mathscr{S}$, see definition after Proposition 3.1.3. Throughout, the manifold topology τ (on $\mathscr{S}$) is assumed to be the Euclidean topology of $\mathscr{S}$, seen as a subset of $\mathbb{R}^N$; that is, the open sets of $\mathscr{S}$ in the manifold topology τ are sets of the form $O \cap \mathscr{S}$, where O is a Euclidean open set of $\mathbb{R}^N$. In Appendix A.1.3 we motivate the choice of such a topology and give further details about this assumption.

Note 2.4.2. Assumption [**SA.1**] will be needed mostly to make sure that all the integral curves of the involved vector fields are well defined (and to guarantee well-posedness of the SDE (1.1)). However see Note 3.1.15 on this point. □

Chapter 3

Properties of UFG diffusions

3.1 Geometrical significance of the UFG condition and implications for the corresponding SDE

In this section we come to explain how the general results outlined in Section 2.3 applies to the study of the dynamics (1.1), assuming that the vector fields $V_0, \ldots, V_d$ satisfy the UFG condition. For clarity, we will compare with the case in which $V_0, \ldots, V_d$ satisfy the Hörmander condition. Subsection 3.1.1 contains *global* results, Subsection 3.1.2 is focussed on *local* results.

3.1.1 Global Results

Recalling the notation and nomenclature of Section 2.2 and motivated by Theorem 2.3.3, we introduce two distributions associated with the vector fields $V_0, \ldots, V_d$; such distributions will play a fundamental role in the analysis of the UFG-dynamics (1.1). Let

- $\hat{\Delta}_0$ be the smallest distribution which contains the space $\text{span}\{V_0, V_1, \ldots, V_d\}$ and is invariant under the vector fields $\{V_0, V_1, \ldots, V_d\}$;
- $\hat{\Delta}$ be the smallest distribution which contains the space $\text{span}\{V_1, \ldots, V_d\}$ and is invariant under the vector fields $\{V_0, V_1, \ldots, V_d\}$.

Let us denote by $n = n(x)$ the rank of the distribution $\hat{\Delta}(x)$. Notice that $n = n(x)$ is a function of the point $x \in \mathbb{R}^N$ and, as such its value can vary from point to point. As Lemma 3.1.1 below demonstrates, if at some point $x \in \mathbb{R}^N$ the rank of $\hat{\Delta}$ is n, then the rank of $\hat{\Delta}_0$ is *at most* $n+1$. We will typically assume that $n < N$, where N is the dimension of the state space $\mathbb{R}^N$ in which the vector fields $V_0, \ldots, V_d$ live, see Note 3.1.12 on this point. We stress that $\hat{\Delta}$ need not contain the vector field V_0 itself (unless for example V_0 is a linear combination of $V_1, \ldots, V_d$). Lemma 3.1.1 below gives a simpler equivalent description of the distributions $\hat{\Delta}$ and $\hat{\Delta}_0$ (which is the one we gave in the introduction).

Lemma 3.1.1. *Let $V_0, \ldots, V_d$ be $d+1$ vector fields on $\mathbb{R}^N$ which satisfy the UFG condition. Recall the decomposition* (1.12)*, the definition of $\mathcal{R}_m$, given in* (2.4)*, and set $\mathcal{R}_{m,0} := \mathcal{R}_m \cup V_0$. Then*

$$\hat{\Delta} = \operatorname{span}\{\mathcal{R}_m\} \quad \textit{and} \quad \hat{\Delta}_0 = \operatorname{span}\{\mathcal{R}_{m,0}\}\,. \tag{3.1}$$

In particular,

$$\hat{\Delta}_0(x) = \operatorname{span}(\hat{\Delta}(x), V_0^{(\perp)}(x)).$$

Proof of Lemma 3.1.1. A proof of this lemma in a general setting can be found in [58, Lemma 1.8.7 and Remark 2.2.3]. For completeness (and to spare the reader from having to compare and match notations and setting with those in [58]), we have added a proof in Appendix A.2.1. □

Note 3.1.2. If the vector fields $V_0, V_1, \ldots, V_d$ satisfies the UFG condition then the distributions $\hat{\Delta}$ and $\hat{\Delta}_0$ are locally of finite type. More precisely, the distributions $\operatorname{span}\{\mathcal{R}_m\}$ and $\operatorname{span}\{\mathcal{R}_m \cup V_0\}$ are *globally* of finite type. This can be checked by using Note 2.2.5 (and the fact that nested commutators can always be expressed as linear combinations of hierarchical commutators, see [51, page 11-12]). □

Since the UFG condition implies that the sets $\mathcal{R}_m$ and $\mathcal{R}_{m,0}$ are locally of finite type, we can apply Theorem 2.3.6 to the distributions given by the span of $\mathcal{R}_m$ and $\mathcal{R}_{m,0}$. By Lemma 3.1.1, the distributions $\hat{\Delta}$ and $\hat{\Delta}_0$ coincide with span of $\mathcal{R}_m$ and $\mathcal{R}_{m,0}$ respectively. As a corollary, we have the following proposition.

Proposition 3.1.3. *If the vector fields $V_0, V_1, \ldots, V_d$ satisfy the UFG condition, then both $\hat{\Delta}_0$ and $\hat{\Delta}$ enjoy the integral manifolds property. In particular the integral manifolds of $\hat{\Delta}_0$ coincide with the orbits of $\hat{\Delta}_0$ (and the same holds for the distribution $\hat{\Delta}$).*

We denote by S ($\mathscr{S}$, respectively) a generic MIM of the distribution $\hat{\Delta}$ ($\hat{\Delta}_0$, respectively). Consistently, S_x ($\mathscr{S}_x$, respectively) will denote the MIM of $\hat{\Delta}$ ($\hat{\Delta}_0$, respectively) through the point $x \in \mathbb{R}^N$. It is easy to see that for every $x \in \mathbb{R}^N$, $S_x \subseteq \mathscr{S}_x$, so that $\mathscr{S}_x$ is a disjoint union of integral manifolds of $\hat{\Delta}$. Notice that $n = n(x)$ is constant along the orbits S of $\hat{\Delta}$.

It is important to observe that any deterministic dynamics started on a maximal integral manifold $\mathscr{S}$ of $\hat{\Delta}_0$ and following the integral curves of the fields $V_0, \ldots, V_d$, will remain in $\mathscr{S}$ for any positive time (see Note 2.3.7). On the other hand, if $X_0 = x$ is the initial datum of the stochastic dynamics (1.1) and $X_0 \in \mathscr{S}_x$, then $X_t \in \overline{\mathscr{S}}_x$ for all $t \geq 0$. This is a consequence of the Stroock and Varadhan support theorem, which we recall below, see [4] for more details.

Theorem 3.1.4 (Stroock and Varadhan)**.** *Let $X_t^{(x)}$ be the solution of the SDE* (1.1)*. The support of the law of $\{X_t^{(x)}\}_{t\in[0,T]}$ in path space, coincides with the closure in $(C([0,T];\mathbb{R}), \|\cdot\|_\infty)$ of the set of paths $(p_t)_{t\in[0,T]}$ such that (p_t) satisfies the control problem:*

$$dp_t = V_0(p_t)dt + \sqrt{2}\sum_{i=1}^{d} V_i(p_t)\psi_i(t)dt, \quad p_0 = x\,, \tag{3.2}$$

for some $\psi_1, , \ldots, \psi_d : [0,T] \to \mathbb{R}$ piecewise constant functions.

Informally, Theorem 3.1.4 says that the stochastic dynamics (1.1) will access in time t the (closure) of the set reachable in time t by the control problem (3.2), as we vary the *controls* $\psi_1, \ldots, \psi_d$ in a suitable set of functions.

Excursus 3.1.5. We would like to further elaborate on the comment started before Theorem 3.1.4. To this end, consider the following one-dimensional SDE:

$$dX_t = -\sin(X_t)dt + \cos(X_t) \circ dB_t\,.^8 \tag{3.3}$$

Here $V_0 = -\sin(x)\partial_x, V_1 = \cos(x)\partial_x$, and $[V_0, V_1] = \partial_x$, so that these fields satisfy

[8] This is a known example, see for example [33].

both the HC and the PHC. According to Chow's theorem (see Note 2.3.7), if V_0, V_1 satisfy the HC then any two points in $\mathbb{R}$ can be joined through integral curves of such fields. However, if we start the dynamics (3.3) at $x \in [-\pi/2, \pi/2]$ then the solution X_t never leaves the interval $[-\pi/2, \pi/2]$. This is not in contradiction to the statement of Chow's theorem. The behaviour of the stochastic dynamics (3.3) is related to the control problem (3.2). On the other hand, when we say that under the HC any two points in $\mathbb{R}^N$ can be joined by integral curves of vectors in $\mathcal{D}$, this is equivalent to saying that the set of points reachable from x by the control system

$$dp_t = V_0(p_t)\psi_0(t)dt + \sum_{i=1}^{d} V_i(p_t)\psi_i(t)dt, \quad p_0 = x, \tag{3.4}$$

is indeed the whole space $\mathbb{R}^N$ (in the above the functions $\psi_1, , \ldots, \psi_d : [0, T] \to \mathbb{R}$ are say piecewise constant controls). Clearly, the set of points accessible by (3.2) is a subset of the set of points accessible by (3.4). In our example, the support of the law of the solution to SDE (3.3) is given by the (closure of the) set of points reachable by the control problem

$$dX_t = -\sin(X_t)dt + \cos(X_t)\psi_1(t)dt.$$

On the other hand, Chow's theorem applied to the vector fields V_0, V_1 refers to the problem

$$dX_t = -\sin(X_t)\psi_0(t)dt + \cos(X_t)\psi_1(t)dt.$$

Such a dynamics can indeed be stirred to access the whole real line, no matter where it is started. □

The theory summarised in Subsection 2.3 describes completely the sets accessible by the control problem (3.4), which are precisely the orbits of the vector fields $V_0, \ldots, V_d$. On the other hand, if we want to study the SDE (1.1) (under the UFG condition) then we are interested in understanding the behaviour of the control problem (3.2). Unfortunately, in full generality, one can only state the following (see [58, Section 2.2]).

Lemma 3.1.6. *With the notation and nomenclature introduced so far, let $V_0, \ldots, V_d$ be smooth vector fields on $\mathbb{R}^N$ satisfying the UFG condition. Then the sets of points* *reachable from x by the control problem* (3.2) *is a subset of $\mathscr{S}_x$ and it contains at least a non-empty open subset of $\mathscr{S}_x$.*

Combining the above and Theorem 3.1.4 we obtain the following.

Proposition 3.1.7. *Consider the SDE* (1.1) *with initial datum $X_0 = x$ and assume that the vector fields $V_0, \ldots, V_d$ satisfy the UFG condition. Then $X_t \in \overline{\mathscr{S}_x}$ for every $t \geq 0$.*[9]

Let us reiterate that Proposition 3.1.7 doesn't say that $X_t^{(x)}$ will explore the whole set $\overline{\mathscr{S}_x}$ (that is, it doesn't imply irreducibility of the process on $\overline{\mathscr{S}_x}$), it simply means that the process X_t will not leave such a set.

Example 3.1.8. For this example we consider the stochastic geodesic equation derived in [28]. The aim of [28] is to study solutions $u(t,x)$ of the stochastic wave geodesic equation on the unit sphere:

$$d\dot{u} = \left(\Delta u + (|\nabla u|^2 - |\dot{u}|^2)u - \frac{1}{2}\dot{u}\right) dt + u \times \dot{u}\, dB_t, \quad |u| = 1, u(0,x) \perp \dot{u}(0,x).$$

Here $\dot{u}$ denotes the time derivative of u, i.e. $\dot{u}(t,x) = \partial_t u(t,x)$. In [28] the authors were concerned with solutions that are independent of the space variables, i.e. $u(t,x) = u(t) = u_t$. By introducing an auxiliary process v_t which is $\mathbb{R}^3$ valued and defined by $v_t := \dot{u}_t$ they find that (u_t, v_t) satisfies the following 6-dimensional Stratonovich SDE:

$$d\begin{pmatrix} u_t \\ v_t \end{pmatrix} = \begin{pmatrix} v_t \\ -|v_t|^2 u_t \end{pmatrix} dt + \begin{pmatrix} 0 \\ u_t \times v_t \end{pmatrix} \circ dB_t. \tag{3.5}$$

In our notation; $X_t = (u_t, v_t)$, $N = 6, d = 1$ and for every $(u,v) \in \mathbb{R}^3 \times \mathbb{R}^3$ we have

$$V_0(u,v) = \begin{pmatrix} v \\ -|v|^2 u \end{pmatrix}, \quad V_1(u,v) = \begin{pmatrix} 0 \\ u \times v \end{pmatrix}.$$

We then have the following commutator relationships:

$$[V_1, V_0] = \begin{pmatrix} u \times v \\ 0 \end{pmatrix}, \; [[V_1, V_0], V_0] = -|v|^2 V_1, \; [[V_1, V_0], V_1] = V_0.$$

[9] We clarify again that the closure is intended to be in the Euclidean topology.

From here we can see that we will generate no new directions by taking further commutators and the distribution $\hat{\Delta}$ satisfies the LFG condition, see Definition 2.3.4; moreover we have

$$\hat{\Delta}_0(u,v) = \hat{\Delta}(u,v) = \mathrm{span}\left(V_1(u,v), [V_1,V_0](u,v), [[V_1,V_0],V_1](u,v)\right).$$

Observe the dimension of $\hat{\Delta}_0$ is 3 except at the point $(0,0)$ where all the vector fields vanish. Define the 3-dimensional manifold

$$M_{r,R} = \{(u,v) : |v| = r, |u| = R, u \perp v\}.$$

Note that the tangent space to $M_{r,R}$ at the point (u,v) is $\hat{\Delta}_0(u,v)$ and $M_{r,R}$ is closed. By Proposition 3.1.7, $X_t \in M_{r,R}$ almost surely when $X_t = (u_t, v_t)$ and $X_0 \in M_{r,R}$. In [28] they consider only the case when $R = 1$, i.e. $|u| = 1$, as they are interested in processes u_t which take values on the unit sphere.

Observe that these vector fields are not globally Lipschitz, however as the solutions always remain in a compact set for fixed initial conditions our results still hold.[10] □

3.1.2 Local considerations: an important change of coordinates

Let $x \in \mathbb{R}^N$ be a *regular point* of a given distribution Δ, i.e. suppose there exists a neighbourhood of x where the dimension of Δ is constant, say equal to n. If this is the case then, locally, there exist n linearly independent vector fields, $\{X_1, \ldots, X_n\} = \mathcal{D}_n$, generating the distribution. Suppose furthermore that $\Delta_{\mathcal{D}_n}$ is involutive and $n < N$ (see Note 3.1.12) . For some small enough $\epsilon > 0$ we can define the map

[10] Indeed, fix some initial conditions u, v with $|u| = 1, |v| = 1$ and construct globally Lipschitz vector fields $\tilde{V}_0, \tilde{V}_1$ with the properties: $\tilde{V}_0(x) = V_0(x)$ and $\tilde{V}_1(x) = V_1(x)$ for any $x \in \mathbb{R}^6$ with $|x| \leq 2$. Let $(\tilde{u}_t, \tilde{v}_t)$ be the solution of

$$d\begin{pmatrix} \tilde{u}_t \\ \tilde{v}_t \end{pmatrix} = \tilde{V}_0(\tilde{u}_t, \tilde{v}_t)dt + \tilde{V}_1(\tilde{u}_t, \tilde{v}_t) \circ dB_t.$$

Then we have that $(\tilde{u}_t, \tilde{v}_t)$ take values in $M_{r,R}$ almost surely and in particular, $|\tilde{u}_t| = 1, |\tilde{v}_t| = r$ for all $t \geq 0$. However since $V_0 = \tilde{V}_0$ and $V_1 = \tilde{V}_1$ on the set $M_{r,R}$ we have that $u_t = \tilde{u}_t$ and $v_t = \tilde{v}_t$ by pathwise uniqueness of solutions to SDEs (see [60, Theorem 5.2.5]). Hence (u_t, v_t) must take values in $M_{r,R}$ almost surely also.

$\Psi : (-\epsilon, \epsilon)^N \to \mathbb{R}^N$ as follows:

$$\begin{aligned} \Psi : (-\epsilon, \epsilon)^N &\longrightarrow \mathbb{R}^N \\ \mathbf{t} := (t_1, \ldots, t_N) &\longrightarrow e^{t_1 X_1} e^{t_2 X_2} \cdots e^{t_N X_N} x \,, \end{aligned}$$

where $X_1, \ldots, X_n$ are as above and $X_{n+1}, \ldots, X_N$ are such that

$$\text{span}\{X_1, \ldots, X_n, X_{n+1}, \ldots, X_N\} = \mathbb{R}^N \quad \text{(at least locally)}.$$

The map Ψ is, at least locally, a diffeomorphism on its image, so it admits an inverse, which we denote by Φ. Differentiating the obvious identity $(\Phi \circ \Psi)(\mathbf{t}) = \mathbf{t}$, one obtains

$$(\mathcal{J}_x \Phi)(\Psi(\mathbf{t})) \cdot (\mathcal{J}_t \Psi)(\mathbf{t}) = \text{Id}_{N \times N}.$$

Let us make the above notation more explicit. The map Φ is a map from (opens sets of) $\mathbb{R}^N$ to (opens sets of) $\mathbb{R}^N$, i.e.

$$\Phi(x) = (\Phi^1(x), \ldots, \Phi^N(x)), \quad x \in \mathbb{R}^N \,,$$

where $\Phi^i : \mathbb{R}^N \to \mathbb{R}$. Therefore the i-th row of the matrix $\mathcal{J}_x \Phi$ is the gradient $\nabla \Phi^i$. On the other hand, the j-th column of the matrix $\mathcal{J}_t \Psi$ is the vector $\frac{\partial \Psi}{\partial t_j} := \{\frac{\partial \Psi^1}{\partial t_j}, \ldots, \frac{\partial \Psi^N}{\partial t_j}\}^T$. The first n columns of the Jacobian matrix $(\mathcal{J}_t \Psi)(\mathbf{t})$ are linearly independent (because Ψ is a diffeomorphism) and, from the above, we have

$$\nabla \Phi^i \cdot \frac{\partial \Psi}{\partial t_j} = 0 \quad \text{for all } j = 1, \ldots, n, i = n+1, \ldots, N. \tag{3.6}$$

By the involutivity of $\Delta_{\mathcal{D}_n}$ the vectors $\{\frac{\partial \Psi}{\partial t_j}\}_{j=1}^n$ belong to $\Delta_{\mathcal{D}_n}$;[11] moreover because they are linearly independent, they span $\Delta_{\mathcal{D}_n}$. Therefore the vectors $\nabla \Phi^i$ are orthogonal to every vector of $\Delta_{\mathcal{D}_n}$, i.e.

$$\nabla \Phi^i \cdot \tau = 0 \quad \text{for every } \tau \in \Delta_{\mathcal{D}_n} \text{ and for every } i = n+1, \ldots, N \,.$$

[11] See e.g. [58, item (ii) on page 25]

Now notice that Φ is (locally) invertible so it can be used as a (local) change of coordinates $\mathbf{z} = \Phi(x)$. With these preliminaries in place, we have the following.

Proposition 3.1.9. *Let Δ be a smooth involutive distribution on $\mathbb{R}^N$ and x_0 a regular point of Δ. In particular, assume that there exists a neighbourhood of x_0 where the dimension of Δ is n. Then there exists a change of coordinates Φ (defined locally) such that*

i) *A vector fields V on $\mathbb{R}^N$ belongs to Δ if and only if in the coordinates defined by Φ, the last $N-n$ components of V are zero;*[12]

ii) *if Δ is invariant under a vector field W then, in the coordinates defined by Φ, the last $N-n$ components of W are functions independent of the first n coordinates. More explicitely, as per notation introduced in (2.1), let*

$$\mathbf{z} = (z^1, \ldots, z^n, z^{n+1}, \ldots, z^N) = (z^1, \ldots, z^N, \zeta, a) = \Phi(x^1, \ldots, x^N)$$

and let $\tilde{W}$ be the representation of W in the new coordinates. Then

$$\tilde{W}(\mathbf{z}) = (\tilde{W}^1(\mathbf{z}), \ldots, \tilde{W}^n(\mathbf{z}), \tilde{W}^{n+1}(z^{n+1}, \ldots, z^N), \ldots, \tilde{W}^N(z^{n+1}, \ldots, z^N))\,.$$

Proof of Proposition 3.1.9. The proof is deferred to Appendix A.2.1. $\square$

We now want to apply Proposition 3.1.9 to the vector fields appearing in the SDE (1.1). We assume that such vector fields on $\mathbb{R}^N$ satisfy the UFG condition at level m. Let $\hat{\Delta}_0$ and $\hat{\Delta}$ be the distributions defined at the beginning of Section 3.1. We know that the rank of $\hat{\Delta}_0$ is constant along the orbits of $\hat{\Delta}_0$ (see comment before Definition 2.3.4). Let $x \in \mathbb{R}^N$ and consider the orbit of $\hat{\Delta}_0$ through x. In view of Lemma 3.1.1, if we assume that $V_0^{(\perp)}(x) \neq 0$ then the rank of $\hat{\Delta}_0$ at x is exactly $n+1$. Recall that N is fixed and it is the dimension of the state space $\mathbb{R}^N$, while $n = n(x)$ is the dimension of the orbits of $\hat{\Delta}$ and it is constant along each one of such orbits. Notice that $\hat{\Delta}$ (and $\hat{\Delta}_0$) is also involutive by construction, so we can use it to apply Proposition 3.1.9.

[12]If $\gamma(t) = e^{tV}x$ and $\tilde{\gamma}(t) = \Phi(\gamma(t))$ then the tangent vector to $\tilde{\gamma}$ is $\tilde{V}(z) = [(\mathcal{J}_\star \Phi) \cdot V(\star)]\,|_{\star=\Phi^{-1}(\mathbf{z})}$.

With this in mind, let us describe the coordinate change. This is obtained by combining the following two steps.

• *Step one*: because $\hat{\Delta}_0 = \text{span}(\mathcal{R}_m, V_0)$ is the tangent space of an $(n+1)$-dimensional submanifold of $\mathbb{R}^N$ one can always *locally* express the vector fields $V_0, \ldots, V_d$ as

$$\tilde{V}_j = (\tilde{V}_j^1, \ldots, \tilde{V}_j^{n+1}, 0, \ldots, 0), \quad j = 0, 1, , \ldots, , d,$$

i.e. the last $N-(n+1)$ coordinates of the vectors $\tilde{V}_j$ are simply zero.

• *Step two*: apply Proposition 3.1.9 using the distribution Δ_n (possibly only to the first $n+1$ coordinates of the involved fields). Then, because $V_1, \ldots, V_d$ belong to Δ_n and V_0 is invariant for Δ_n, one obtains, in the new local coordinates, (and recalling the notation introduced in Section 2.1)

$$\tilde{V}_0 = (\tilde{V}_0^1(\mathbf{z}), \ldots, \tilde{V}_0^n(\mathbf{z}), \tilde{V}_0^{n+1}(\zeta, a), 0, \ldots, 0)$$

$$\tilde{V}_j = (\tilde{V}_j^1(\mathbf{z}), \ldots, \tilde{V}_j^n(\mathbf{z}), 0, \ldots, 0), \quad j = 1, \ldots, d\,,$$

where we keep the same notation $\tilde{V}_j$ for the new representation of the vector fields after this further change of coordinates. This shows that, in the new coordinates, the vector fields $V_0, \ldots, V_d$ take the form (2.2) - (2.3).

We now want to express the SDE (1.1) in the new local coordinates. If X_t is the original process, $\mathbf{Z}_t$ is the process in the new coordinates. In particular

$$\mathbf{Z}_t = (Z_t, \zeta_t, a_t),$$

where $Z_t \in \mathbb{R}^n$ contains the first n coordinates of $\mathbf{Z}_t$, ζ_t is the $(n+1)$-th coordinate of the process and a contains the remaining $N-(n+1)$ components (which do not change in time, see below). Putting everything together and using the convention (2.2) - (2.3), one obtains that, in the new coordinates, the SDE (1.1) with initial datum $\mathbf{Z}_0 = (z_0, \zeta_0, a_0)$ is simply

$$Z_t = z_0 + \int_0^t U_0(Z_s, \zeta_s, a_0)\, ds + \sum_{j=1}^{d} \int_0^t U_j(Z_s, \zeta_s, a_0) \circ dB_s^j \tag{3.7}$$

$$\zeta_t = \zeta_0 + \int_0^t W_0(\zeta_s, a_0)\, ds \tag{3.8}$$

$$a_t = a_0\,. \tag{3.9}$$

Notice that from the above one can also deduce that, in the new coordinates, $\tilde{V_0}^{(\hat{\Delta})} = (U_0, 0, \ldots, 0)$ while $\tilde{V_0}^{(\perp)} = (0, \ldots, 0, W_0, 0, \ldots, 0)$. Assuming for the moment that at the initial point $x = X_0$ the dimension of $\hat{\Delta}_0$ is exactly $n+1$, the fact that the last $N-(n+1)$ components of the dynamics remain constant reflects the fact that, at least for a short enough time, the solution of the SDE remains in the integral submanifold of $\hat{\Delta}_0$ from which it started, coherently with Lemma 3.1.6 and Proposition 3.1.7.

If at the initial point the rank of $\hat{\Delta}_0$ is exactly N, i.e. $n + 1 = N$, then one simply has

$$Z_t = z_0 + \int_0^t U_0(Z_s, \zeta_s)\, ds + \sum_{j=1}^{d} \int_0^t U_j(Z_s, \zeta_s) \circ dB_s^j \tag{3.10}$$

$$\zeta_t = \zeta_0 + \int_0^t W_0(\zeta_s)\, ds\,, \tag{3.11}$$

and this time $\tilde{V_0}^{(\hat{\Delta})} = (U_0, 0, \ldots, 0)$ while $\tilde{V_0}^{(\perp)} = (0, \ldots, 0, W_0)$. In this simpler case it is clearer that we have locally reduced the SDE (1.1) to an ODE component, ζ_t (which evolves independently of all the other components) and an $(N-1)$ - dimensional SDE. We emphasize that, because the change of coordinates is local, such a representation will hold only for small enough t.

Example 3.1.10 (UFG-Heisenberg)**.** Consider the following dynamics in $\mathbb{R}^3$

$$dX_t = -X_t dt$$

$$dY_t = -Y_t dt + \sqrt{2} dW_t^2$$

$$dZ_t = -2Z_t dt - \sqrt{2} Y_t \circ dW_t^1 + \sqrt{2} X_t \circ dW_t^2$$

Here $V_0 = (-x, -y, -2z)$, $V_1 = (0, 0, -y)$, $V_2 = (0, 1, x)$. This example was introduced in [32] and named the *UFG-Heisenberg dynamics* (as it comes from a modification of the Heisenberg group). This is already globally in the form ODE+SDE. The ODE for the first coordinate can be solved explicitly, giving $X_t = x_0 e^{-t}$. Therefore, if we start the dynamics at (x_0, y_0, z_0) with $x_0 > 0$ ($x_0 < 0$, respectively), then the system evolves (at least for finite time) in the semispace with positive x-coordinates (negative, respectively). If the initial datum is on the plane $(0, y_0, z_0)$ then the dynamics remains confined to such a plane for all subsequent times. This is coherent with the following: for the above set of vector fields, one has $\hat{\Delta}_0((x, y, z)) \simeq \mathbb{R}^3$ if $x > 0$ or $x < 0$ and $\hat{\Delta}_0((x, y, z)) \simeq \mathbb{R}^2$ when $x = 0$. The distribution $\hat{\Delta}_0$ has three orbits, namely the sets

$$\mathscr{S}_+ = \{(x, y, z) \in \mathbb{R}^3 : x > 0\}, \quad \mathscr{S}_- = \{(x, y, z) \in \mathbb{R}^3 : x < 0\},$$

$$\mathscr{S}_0 = \{(x, y, z) \in \mathbb{R}^3 : x = 0\}.$$

As for the distribution $\hat{\Delta}$, this spans $\mathbb{R}^2$ at every point. Moreover, the orbit of $\hat{\Delta}$ through the point (b, y, z) is the plane $x = b$. For this reason, when working on this example we will simply denote by S_b the orbit through the point (b, y, z). In particular, notice that $\mathscr{S}_0 = S_0$. □

Example 3.1.11 (Random Circles)**.** Consider the SDE

$$dX_t = -Y_t dt + \sqrt{2} X_t \circ dB_t \tag{3.12}$$

$$dY_t = X_t dt + \sqrt{2} Y_t \circ dB_t, \tag{3.13}$$

where B_t is a one-dimensional Brownian Motion. This system satisfies neither the HC nor the PHC, however the UFG condition is satisfied at level $m = 1$. Indeed we have

$$V_0(x,y) = \begin{pmatrix} -y \\ x \end{pmatrix}, \quad V_1(x,y) = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{and} \quad [V_1, V_0] = 0\,.$$

For every $(x,y) \in \mathbb{R}^2$, $\hat{\Delta}(x,y) = \operatorname{span}\{V_1(x,y)\}$; except for the origin, the orbits of $\hat{\Delta}$ are radial half-lines. That is, $S_{(x,y)} = (0,0)$ if $(x,y) = (0,0)$ and $S_{(x,y)} = \{(sx, sy), s > 0\}$ otherwise. Indeed, $S_{(x,y)}$ coincides with the set of points accessible by the integral curves of V_1, which can be found explicitly:

$$e^{tV_1}(x,y) = \begin{pmatrix} x\, e^t \\ y\, e^t \end{pmatrix}, \quad t \in \mathbb{R}\,.$$

Moreover, V_0 is orthogonal to V_1, so $V_0^{(\hat{\Delta})} = 0$ and $V_0^{(\perp)} = V_0$; therefore $\hat{\Delta}_0(0,0) = \{(0,0)\}$, $\hat{\Delta}_0(x,y) = \mathbb{R}^2$ outside the origin, $\mathscr{S}_{(x,y)} = \mathbb{R}^2 \setminus \{(0,0)\}$ if $(x,y) \neq (0,0)$ and $\mathscr{S}_{(0,0)} = \{(0,0)\}$. In this example the local change of coordinates in the neighbourhood of $(1,0)$ is given by the diffeomorphism

$$\Phi_{(1,0)}(x,y) = \left(\arctan\left(\frac{y}{x}\right), \frac{1}{2}\log(x^2+y^2)\right).$$

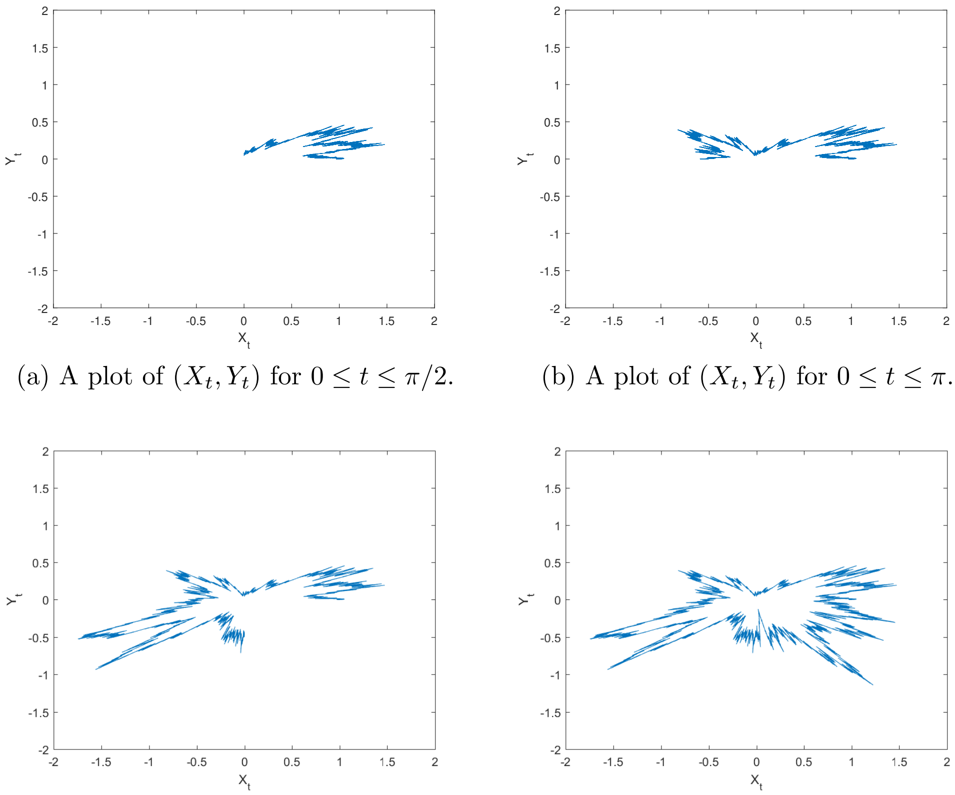
After such a change of coordinates, the SDE (3.12) - (3.13) can be expressed, locally, as

$$d\zeta_t = dt \tag{3.14}$$

$$dZ_t = \sqrt{2}dW_t \tag{3.15}$$

Let $C_t = (X_t, Y_t) \in \mathbb{R}^2$. In Figure 3.1 below we plot the evolution of C_t, i.e. the solution of (3.12) - (3.13). From the plots it should be clear that (ζ_t, Z_t) are just the polar coordinates of the point C_t: ζ_t represents the angle, which evolves deterministically with a simple anticlockwise motion, while Z_t (or, to be more precise, $\exp(2Z_t)$) is the radius, which changes randomly according to the SDE (3.15). □

Note 3.1.12. If the dimension n of $\hat{\Delta}$ was equal to N for every $x \in \mathbb{R}^N$, this would



(a) A plot of (X_t, Y_t) for $0 \leq t \leq \pi/2$.

(b) A plot of (X_t, Y_t) for $0 \leq t \leq \pi$.

(c) A plot of (X_t, Y_t) for $0 \leq t \leq \frac{3\pi}{2}$

(d) A plot of (X_t, Y_t) for $0 \leq t \leq 2\pi$

Figure 3.1: The process (X_t, Y_t) of Example 3.1.11, with initial condition $(X_0, Y_0) = (1, 0)$. The angle of rotation evolves deterministically in an anticlockwise direction, while the radius changes randomly, according to (3.15).

imply that $\hat{\Delta}(x) = \hat{\Delta}_0(x)$ for every $x \in \mathbb{R}^N$. In particular, the Parabolic Hörmander Condition (**PHC**) would hold. This case is well studied in the literature and we do not wish to consider it here. For this reason many of the statements of this section are made under the assumption that $n < N$. We need to emphasize that it may happen that the two distributions coincide on a manifold (see Example 3.1.10, where the two distributions coincide on the plane $x = 0$) and it may also happen that they both have full rank N on a manifold, while they differ on other manifolds (see Example 3.1.14 below). The case that is not interesting to our purposes is the one in which they coincide and have full rank on the whole of $\mathbb{R}^N$. Most of our theorems do cover that case as well (unless otherwise explicitly stated); but they are not really conceived in that framework. $\square$

Note 3.1.13. The change of coordinates illustrated in this section will be an important technical tool throughout. We would like to point out how such a change of coordinates gives a different (and complementary) perspective on the smoothness results of Kusuoka and Stroock and of Crisan et al [21–24, 30] that we mentioned in the Introduction. As recalled in Section 1.1, in these works the authors show that if f is a continuous and bounded function then, under the UFG condition, the function $(\mathcal{P}_t f)(x)$ is not necessarily smooth in every direction (as it would be the case under the Hörmander condition), but it is in general only smooth in the directions $V_{[\alpha]}$, $\alpha \in \mathcal{A}_m$. In particular, it may not be differentiable in the direction V_0. In view of the decomposition (1.12) and of the change of coordinates presented in this section, this result is quite intuitive, as we explain. By (1.12), it is clear that if $V_0^{(\perp)} = 0$ then $(\mathcal{P}_t f)(x)$ is differentiable in the direction V_0 (as in this case V_0 is a combination of the vectors in $\mathcal{R}_m$) and, as a consequence, it is differentiable in t as well. The loss of smoothness happens if and only if $V_0^{(\perp)} \neq 0$. For simplicity (and without any loss of generality), let us restrict to a manifold where $n+1 = N$, so that the local change of coordinates gives (3.10)-(3.11). As already observed, the representation of $V_0^{(\perp)}$ in the new coordinates is given by $\tilde{V_0}^{(\perp)} = (0, \ldots, 0, W_0)$, where W_0 is the function driving the ODE component. Hence $V_0^{(\perp)}$ is inherently linked to the deterministic part of the system, which clearly doesn't provide any smoothness. This also explains why, while there is no smoothness in the direction V_0, the semigroup will always be smooth in the direction $\partial_t - V_0$ (to be more precise, in the direction $\partial_t - V_0^{(\perp)}$), as solutions of the ODE are constant in this direction. Finally, the deterministic part of the dynamics is responsible for the lack of density (i.e. for the fact that the law of the process does not admit a density on $\mathbb{R}^N$). It is useful to the purposes of this discussion to point out that the one-dimensional transport equation is an extreme example of UFG condition; that is, consider the PDE $\partial_t u(t,x) = \partial_x u(t,x), (t,x) \in \mathbb{R}_+ \times \mathbb{R}^N$, with initial datum $u(0,x) = f(x)$. Here $V_0 = \partial_x$. As is well known, the solution to such a PDE is just $u(t,x) = f(x+t)$, hence no smoothing occurs in the space direction. However the solution is smooth in the direction $(\partial_t - \partial_x) = \partial_t - V_0$, as it is constant in such a direction. Therefore, UFG diffusions include a vast range of behaviours, from smooth elliptic diffusions to deterministic equations. □

Example 3.1.14. In $\mathbb{R}^2$ let $V_0 = 1_A \partial_x$ and $V_1 = 1_{A^c}\partial_x + 1_{A^c}\partial_y$ (strictly speaking here the coefficients are not smooth), where A is the set $A = \{(x,y) \in \mathbb{R} : x \in [-1,1]\}$. Then $\hat{\Delta}(x,y) = \hat{\Delta}_0(x,y)$ and they are both two-dimensional for every

$(x,y) \in A^c$ while $\hat{\Delta}(x,y) \neq \hat{\Delta}_0(x,y)$ for every $(x,y) \in A$, as on this set $\hat{\Delta}_0$ is one dimensional while $\hat{\Delta}(x,y) = 0$ for every $(x,y) \in A$. $\square$

Note 3.1.15. A final note on a technical point: as we have emphasized, to avoid having problems with the well-posedness of the integral curves, we work under the standing assumption [**SA.1**]. After the change of coordinates the coefficients of the vector fields (in the new coordinates) may grow more than linearly, but they will still be smooth. Hence, in the neighbourhood in which they are defined, the vector fields will still be locally Lipschitz. The situation is more delicate with the vector $V_0^{(\perp)}$: if V_0 is smooth, this is not the case for $V_0^{(\perp)}$ as well, see Example 5.3.9. Whenever this may cause issues, we will assume that $V_0^{(\perp)}$ is at least such that the integral curve of $V_0^{(\perp)}$ through a given point is unique and well defined (at least on given manifolds). $\square$

We conclude this section by stating a couple of technical lemmata which will be useful in the following.

Lemma 3.1.16. *Assume the vector fields $V_0, \ldots, V_d$ satisfy the UFG condition. Let $\mathscr{S}$ be a maximal integral manifold of $\hat{\Delta}_0$ and S be an integral submanifold of $\hat{\Delta}$ such that $S \subseteq \mathscr{S}$. Then $\partial S := \bar{S} \setminus S$ is contained within $\partial \mathscr{S} := \bar{\mathscr{S}} \setminus \mathscr{S}$.*[13]

Proof of Lemma 3.1.16. The proof is deferred to Appendix A.2.1. $\square$

The statement of Lemma 3.1.16 would clearly not be true if S and $\mathscr{S}$ were two arbitrary sets, it only holds because of the particular structure of the integral manifolds of $\hat{\Delta}$ and $\hat{\Delta}_0$. As a side remark, notice that while $S \subseteq \mathscr{S}$ implies $\partial S \subseteq \partial \mathscr{S}$, it is not the case, in general, that the boundary of $\mathscr{S}$ is the union of boundaries of orbits of $\hat{\Delta}$, see Example 3.1.10.

Lemma 3.1.17. *With the notation introduced so far, assume the vector fields $V_0, \ldots, V_d$ satisfy the UFG condition. Let $x_0 \in \mathbb{R}^N$ and recall that x_0 belongs to exactly one integral manifold of $\hat{\Delta}_0$, the manifold $\mathscr{S}_{x_0}$. Consider the vector field $V_0^{(\perp)}$ (defined in* (1.12)*) and assume such a vector field is smooth. Then either $V_0^{(\perp)}(x) = 0$ for every $x \in \mathscr{S}_x$ or $V_0^{(\perp)}(x) \neq 0$ for every $x \in \mathscr{S}_x$.*

Proof of Lemma 3.1.17. The proof is deferred to Appendix A.2.1. $\square$

[13]Closures are meant in the Euclidean topology, see Appendix A.1.3.

3.2 Qualitative Results on UFG diffusions

In this section we study the behaviour of the diffusion X_t (1.1) under the sole assumption that the vector fields $V_0, \dots, V_d$ appearing in (1.1) satisfy the UFG condition. As observed also in [32, Note 4.3], under the sole UFG condition one cannot expect to make any quantitative deductions on the behaviour of the process X_t. Neither can one expect the UFG condition itself to imply any results about existstence or uniqueness of invariant measures, as there are many elliptic diffusions that don't have an invariant measure (the simplest example being Brownian motion on $\mathbb{R}$). In order to study invariant measures and decay to equilibrium we will have to make further assumptions. Nonetheless, the geometric considerations made in the previous sections allow us to prove several qualitative statements on the behaviour of the diffusion. The main results of this section are Proposition 3.2.1, Proposition 3.2.3 and Proposition 3.2.7. Collectively, these three results impart a lot of intuition about UFG dynamics and cointain a lot of useful information. After each one of these three statements we have inserted a note to comment on the meaning of these propositions, see Note 3.2.2, Note 3.2.4 and Note 3.2.8. The results of Section 5.1 and Section 5.2 heavily rely on the statements of this section.

Recall that we denote by S ($\mathscr{S}$, respectively) a generic integral manifold of the distribution $\hat{\Delta}$ ($\hat{\Delta}_0$, respectively). Consistently, S_x ($\mathscr{S}_x$, respectively) denote the integral manifold of $\hat{\Delta}$ ($\hat{\Delta}_0$, respectively) through the point $x \in \mathbb{R}^N$.

Proposition 3.2.1. *Assume that the vector fields $V_0, V_1, \dots, V_d$ satisfy the UFG condition and let X_t be the solution of the SDE* (1.1). *Let $\mathscr{S}$ be a maximal integral manifold of $\hat{\Delta}_0$ and let $\partial\mathscr{S}$ be the boundary of $\mathscr{S}$, i.e. $\partial\mathscr{S} := \bar{\mathscr{S}} \setminus \mathscr{S}$. Then the following holds:*

i) *If $\partial\mathscr{S}$ is not empty, it is a union of integral submanifolds of $\hat{\Delta}_0$;*

ii) *If $X_0 = x \in \partial\mathscr{S}$ then $X_t \in \partial\mathscr{S}$ for all $t > 0$ (almost surely).*

Proof of Proposition 3.2.1. The proof is deferred to Appendix A.2.1. $\square$

Note 3.2.2. Let us explain the meaning and consequences of Proposition 3.2.1. Suppose we start the SDE (1.1) at $x \in \mathbb{R}^N$. Because the integral manifolds of $\hat{\Delta}_0$ partition $\mathbb{R}^N$, x belongs to one of such integral manifolds, the one which we denote by $\mathscr{S}_x$. As a consequence of Proposition 3.1.7 we know that the process will never leave the closure of $\mathscr{S}_x$; however, if it started in the interior, it could in principle hit the boundary (which is a manifold whose dimension is lower than the dimension of $\mathscr{S}_x$) and then come back to the interior. What we prove here is that this is not possible. Furthermore, because the boundary of $\mathscr{S}_x$ is itself a union of integral manifolds of $\hat{\Delta}_0$, one could repeat the previous reasoning once the process enters the boundary (if this is the case). As a result of iterating this line of thought, we have that, along the path of $X_t^{(x)}$, the rank of the distribution $\hat{\Delta}_0$ can only decrease (or stay the same). In other words, we have shown that for every $x \in \mathbb{R}^N$ and $t \leq u$, one has

$$\operatorname{rank}(\hat{\Delta}_0(X_u^{(x)})) \leq \operatorname{rank}(\hat{\Delta}_0(X_t^{(x)})).$$

$\square$

Before stating the next result we recall that the vector $V_0^{(\perp)}$ has been defined in (1.12).

Proposition 3.2.3. *Let X_t be the solution of the SDE* (1.1) *with initial condition $x_0 \in \mathbb{R}^N$. If the vector fields $V_0, V_1, \ldots, V_d$ appearing in* (1.1) *satisfy the UFG condition then*

$$X_t^{(x_0)} \in \overline{S}_{e^{tV_0^{(\perp)}}(x_0)}, \quad \textit{almost surely.}$$

We clarify that $\overline{S}_{e^{tV_0^{(\perp)}}(x_0)}$ is the closure (in the Euclidean topology) of the integral manifold of $\hat{\Delta}$ through the point $e^{tV_0^{(\perp)}}(x_0) \in \mathbb{R}^N$.

Proof of Proposition 3.2.3. If $V_0^{(\perp)}(x_0) = 0$ then the result follows immediately from Proposition 3.1.7 and Lemma 3.1.1. Indeed, by Proposition 3.1.7 we know that $X_t^{(x_0)} \in \bar{\mathscr{S}}_{x_0}$ and by Lemma 3.1.1 (and Lemma 3.1.17) we have $\mathscr{S}_{x_0} = S_{x_0}$. So we only need to treat the case $V_0^{(\perp)}(x_0) \neq 0$. This will be done by considering the control problem associated with the SDE (1.1) and by using Stroock and Varadhan Support Theorem. We postpone this part of the proof to Appendix A.2.1. $\square$

Note 3.2.4. Proposition 3.2.3 clarifies the pivotal role of the vector $V_0^{(\perp)}$. To convey more intuition about the role of $V_0^{(\perp)}$, let us assume that $V_0^{(\perp)}(x) \neq 0$ for every x in $\mathscr{S}_{x_0}$, x_0 being the starting point of the SDE (1.1). We already know by Proposition 3.1.7 that $X_t^{(x_0)}$ will not leave $\bar{\mathscr{S}}_{x_0}$, so that we can consider $\bar{\mathscr{S}}_{x_0}$ to be the state space of the dynamics. As already observed before Proposition 3.1.3, every $x \in \mathscr{S}_{x_0}$, belongs to exactly one orbit S of $\hat{\Delta}$ and, moreover, the union of the manifolds $\{S_x\}_{x\in\mathscr{S}_{x_0}}$ gives precisely $\mathscr{S}_{x_0}$. In other words, the orbits of $\hat{\Delta}$ that belong to $\mathscr{S}_{x_0}$ partition $\mathscr{S}_{x_0}$. Furthermore, because $V_0^{(\perp)} \neq 0$ on $\mathscr{S}_{x_0}$ and the rank of $\hat{\Delta}_0$ is constant on $\mathscr{S}_{x_0}$, one has (see Lemma 3.1.1) that if $\mathscr{S}_{x_0}$ has rank $n+1$ then every orbit $S_x, x \in \mathscr{S}_{x_0}$, will be a manifold of dimension n. In particular, there is no $x \in \mathscr{S}_{x_0}$ such that $S_x = \mathscr{S}_x$ (so that the partition of $\mathscr{S}_{x_0}$ into orbits of the distribution $\hat{\Delta}$ is not the trivial one). With this premise, it makes sense to ask the following question: while we know that the process will not leave $\overline{\mathscr{S}}_{x_0}$ for every $t \geq 0$, if we fix an arbitrary positive time $t > 0$, can we tell more precisely where, within $\mathscr{S}_{x_0}$, $X_t^{(x_0)}$ is? In particular, can we determine which submanifold S it belongs to, i.e. which element of the partition of $\mathscr{S}_{x_0}$ is visited at time $t \geq 0$? The answer, given by Proposition 3.2.3, is the following: let $y = e^{tV_0^{(\perp)}} x_0$. Then, while $x_0 \in S_{x_0}$, $X_t \in \overline{S}_y$. In other words, the vector $V_0^{(\perp)}$ will make the SDE move from one submanifold of the partition (of $\mathscr{S}_{x_0}$) to another. Another question is whether it is possible that X_t will visit one of such submanifolds twice or whether it is the case that, once one of these submanifolds has been visited, it will never be hit again. Example 3.2.6 below shows that the submanifolds of the partition can be visited an arbitrary number of times. □

Example 3.2.5. Recall the UFG-Heisenberg SDE introduced in Example 3.1.10. In this case $V_0^{(\perp)} = (-x, 0, 0)$ and, as we have already mentioned, $S_{(x_0,y_0,z_0)}$ is the plane $S_{(x_0,y_0,z_0)} = \{(x, y, z) : x = x_0\}$. If $V_0^{(\perp)} = (-x, 0, 0)$ then the integral curve of $V_0^{(\perp)}$ through (x_0, y_0, z_0) is $e^{tV_0^{(\perp)}}(x_0, y_0, z_0) = (e^{-t}x_0, y_0, z_0)$ so that

$$S_{e^{tV_0^{(\perp)}}(x_0,y_0,z_0)} = \{(x, y, z) \in \mathbb{R}^3 : x = e^{-t}x_0\}.$$

It is therefore clear that if $(X_0, Y_0, Z_0) = (x_0, y_0, z_0)$ then $(X_t, Y_t, Z_t) = (x_0 e^{-t}, Y_t, Z_t) \in \overline{S}_{(e^{-t}x_0,y_0,z_0)}$. □

Example 3.2.6 (Random Circles, Example 3.1.11, continued)**.** Let us go back to

Example 3.1.11. Consider the integral curve of $V_0^{(\perp)}$, namely

$$e^{tV_0}(x,y) = e^{tV_0^{(\perp)}}(x,y) = \begin{pmatrix} x\cos(t) - y\sin(t) \\ x\sin(t) + y\cos(t) \end{pmatrix}. \tag{3.16}$$

To fix ideas, let $(x_0, y_0) = (1,0)$ be the initial condition of the SDE; then the integral curve of $V_0^{(\perp)}$ through $(x_0, y_0) = (1,0)$ is the unit circle:

$$e^{tV_0^{(\perp)}}(1,0) = (\cos(t), \sin(t))$$

and $S_{e^{tV_0^{(\perp)}}(1,0)} = S_{(\cos(t),\sin(t))}$ is the (open half) radial line at an angle t from the x-axis; that is, it is the (open half) radial line that intersects the unit circle at the point $(\cos(t), \sin(t))$. On the other hand the solution of the SDE with initial datum $(x_0, y_0) = (1,0)$ is given by

$$X_t^{(x_0,y_0)} = \cos(t)e^{\sqrt{2}B_t}, \tag{3.17}$$

$$Y_t^{(x_0,y_0)} = \sin(t)e^{\sqrt{2}B_t}. \tag{3.18}$$

Therefore one can again explicitly verify that for every $t > 0$, $(X_t^{(x_0,y_0)}, Y_t^{(x_0,y_0)})$ belongs to $\overline{S}_{(\cos(t),\sin(t))}$. □

Proposition 3.2.7. *With the notation introduced so far, assume the vector fields $V_0, \ldots, V_d$ satisfy the UFG condition and let*

$$E_t := \{x \in \mathbb{R}^N : \mathbb{P}_x(X_t^{(x)} \notin \mathscr{S}_x) > 0\}$$

and

$$E := \{x \in \mathbb{R}^N : \mathbb{P}_x(X_t^{(x)} \notin \mathscr{S}_x) > 0 \text{ for some } t > 0\}.$$

Then, for any invariant measure μ of the SDE (1.1) *(should at least one exist), we have $\mu(E_t) = 0$ for every $t > 0$. As a consequence, $\mu(E) = 0$ as well.*

Proof of Proposition 3.2.7. The proof is deferred to Appendix A.2.1. □

Note 3.2.8. Informally, Proposition 3.2.7 says that any invariant measure (should at least one exist) gives zero weight to the set of points that, under the action of the dynamics prescribed by the SDE (1.1), leave in finite time the submanifold from which they start. That is, the set of points x such that $X_t^{(x)} \notin \mathscr{S}_x$ for some time $t > 0$, has μ-measure zero. In view of Proposition 3.2.1 this result is intuitive: in general, if the dynamics leaves a set it can return infinitely many times to that set (when this happens the set is said to be recurrent). Because along the trajectories of $X_t^{(x)}$ the rank of the distribution can only decrease, if the process $X_t^{(x)}$ leaves the integral manifold $\mathscr{S}_x$ from which it started, it will never return to it. The dynamics will therefore spend an infinite amount of time outside the manifold $\mathscr{S}_x$, so that the invariant measure, if it exists, it can only be supported outside such a manifold. In other words, the theorem says that an integral submanifold $\mathscr{S}$ is a recurrent set if and only if the process never leaves it (once it enters it). This argument constitutes an informal proof of the theorem. Notice also that this theorem doesn't say anything about say Geometric Brownian motion (see Example 2.2.3) or the UFG-Heisenberg process of Example 3.1.10, as such dynamics only leave the initial submanifold in infinite time; for any finite time they stay in the submanifold from which they started. □

Lemma 3.2.9. *Assume the vector fields $V_0, \ldots, V_d$ appearing in* (1.1) *satisfy the UFG condition and that the long-time derivative estimate* (2.9) *holds for the semi-group $\mathcal{P}_t$ defined in* (1.2)*. Then, given a maximal integral submanifold S of $\hat{\Delta}$, we have*

$$\lim_{t\to\infty} |\mathcal{P}_t f(x) - \mathcal{P}_t f(y)| = 0, \tag{3.19}$$

for all $f \in C_b(\mathbb{R}^N)$ and $x, y \in S$.

Proof of Lemma 3.2.9. The proof is deferred to Appendix A.2.1 □

Proposition 3.2.10. *Consider the assumptions and setting of the previous lemma and let S be a maximal integral manifold of $\hat{\Delta}$. Then, among all the invariant measures μ of* (1.1) *(assuming at least one such measure exists), there exists at most one such that $\mu(S) = 1$. Moreover, if such a measure exists, then it is ergodic (in the sense that $\mathcal{P}_t \mathbb{1}_E = \mathbb{1}_E$ for some Borel set E, implies that $\mu(E) = 1$ or 0) and for every $x \in S$ and $f \in C_b(\mathbb{R}^N)$ we have*

$$\mathcal{P}_t f(x) \to \int_S f(y)\mu(dy). \tag{3.20}$$

Proof of Proposition 3.2.10. The proof is deferred to Appendix A.2.1. □

Chapter 4

Pathwise Approach to Derivative Estimates

4.1 A pathwise version of the Bakry-Emery approach to derivative estimates for Markov semigroups

In this section and the next we study derivative estimates for Markov semigroups, i.e. we study sufficient conditions in order for bounds of the type (1.15) to hold. To be more precise, in this section we find conditions in order for (1.18) (or better, (4.6)) to hold, in Section 4.2 we will give criteria to obtain (1.15) from (1.18). To begin with, let us clarify the setting in which we will work.

For this chapter we shall use the Itô form of the SDE (1.1), as the vector fields are smooth we may write (1.1) in Itô form as

$$X_t = X_0 + \int_0^t V_{0,\text{Itô}}(X_s)ds + \sqrt{2}\sum_{i=1}^{d}\int_0^t V_i(X_s)dB^i(s), X_0 = x, \tag{4.1}$$

where

$$V^i_{0,\text{Itô}}(x) = V^i_0(x) + \sum_{k=1}^{d}\sum_{j=1}^{N} V^j_k(x)\partial_j V^i_k(x). \tag{4.2}$$

In this section we will consider SDEs of the form (1.1) and restrict to the case $N = 1$. Without loss of generality, by Lemma A.2.2, we may assume that $d = 1$ as well and consider one-dimensional SDEs of the form

$$dX_t^{(x)} = V_0(X_t^{(x)}) + \sqrt{2}V_1(X_t^{(x)}) \circ dB_t, \quad X_0^{(x)} = x. \tag{4.3}$$

Note 4.1.1. As we recall in Lemma A.2.3, any one-dimensional SDE with multiplicative noise can be transformed into a (one-dimensional) SDE with additive noise (i.e into an SDE of the form (A.24)). After such a transformation the differential operator V_1 is therefore just the derivative in direction x, $V_1 = \partial_x$. Hence, in the elliptic case, one can always recover derivative estimates in the coordinate direction ∂_x from derivative estimates in the direction V_1. □

We shall concentrate on estimates for first order derivatives however similar arguments could be applied to higher order derivatives as shall be demonstrated in Lemma 4.1.10 for a class of examples.

While we assume globally Lipschitz coefficients (see **[SA.1]**), the case we really have in mind in developing in this section and the next is the one in which the coefficients of the SDE are bounded (as well as Lipshitz). To explain why this case is harder then when one has linear growth of the coefficients, let us start by recalling that in [32] the authors proved that, under the OAC (1.16), the estimate (1.15) holds; however, as we show in Lemma A.2.1 for a large class of SDEs the OAC implies unboundedness of the coefficients of the SDE. On the other hand, one does expect that exponential decay of derivatives of the semigroup may hold even if the coefficients of the SDE are bounded. To illustrate why this is the case on a (relatively) simple example, start by considering the one-dimensional ODE

$$\frac{d}{dt}\xi_t^{(x)} = -\arctan(\xi_t^{(x)}), \quad \xi_0 = x\,.$$

This ODE has a single equilibrium at $\xi = 0$ and such an equilibrium is stable.

Moreover, for any $x \in \mathbb{R}$, we have[14]

$$\partial_x(\xi_t^{(x)}) \leq \exp\left(-\frac{t}{1+x^2}\right). \tag{4.4}$$

Motivated by this analogy we shall consider the SDE

$$dX_t^{(x)} = -\arctan(X_t^{(x)})dt + \sqrt{2}dB_t. \tag{4.5}$$

In Example 4.1.7 and Example 4.3.3 we will show that (1.15) does hold for the above SDE. Although this example *does not* satisfy the OAC (1.16), one can easily verify that for each $R > 0$ and f sufficiently smooth we have

$$([V_1, V_0]f)(x)(V_1 f)(x) \leq -\frac{1}{1+R^2}|(V_1 f)(x)|^2, \quad \text{for every } x \in [-R, R].$$

That is, the OAC is locally satisfied for $x \in [-R, R]$; this motivates us to introduce local versions (1.17) of the OAC.

Note 4.1.2. We note in passing that the solution of (4.5) has uniformly in time bounded exponential moments, i.e.

$$\sup_{t \geq 0} \mathbb{E}[e^{|X_t^x|}] < \infty, \quad \forall x \in \mathbb{R}.$$

So, overall, on any fixed interval we have a version of the Obtuse Angle Condition and the probability of the process leaving an interval is exponentially small (for each $R > 0$ the probability $X_t \notin [-R, R]$ is bounded by Ce^{-R} by Markov's inequality). $\square$

Because of the local nature of (1.17), in this section we shall develop a pathwise approach to obtaining exponential decay (1.15) of the derivative in direction V_1 of

[14] Indeed, differentiating (4.1) with respect to x gives

$$\frac{d}{dt}\partial_x \xi_t^{(x)} = -\frac{1}{1+(\xi_t^{(x)})^2}\partial_x \xi_t^{(x)}.$$

We can solve this to find

$$\partial_x \xi_t^{(x)} = \exp\left(-\int_0^t \frac{1}{1+(\xi_s^{(x)})^2} ds\right).$$

Finally, since $\xi_s^{(x)}$ converges monotonically towards zero we have $(\xi_s^{(x)})^2 \leq x^2$ and hence (4.4) follows.

the semigroup under the condition (1.17). Note that, even in this one-dimensional setting, since we are not assuming ellipticity of equation (4.3), it is not obvious that the only vector field in which we should be interested is V_1. In Lemma A.2.4 we clarify why this is the case.

We now move on to proving that if the LOAC (1.17) is satisfied then, for every $t \geq 0, x \in \mathbb{R}, f \in C_{V_1}(\mathbb{R})$

$$|V_1 \mathcal{P}_t f(x)| \leq \mathbb{E}\left[\exp\left(-2\int_0^t \lambda(X_r^{(x)})dr\right)\right]^{\frac{1}{2}} \|V_1 f\|_\infty. \tag{4.6}$$

Here $C_{V_1}(\mathbb{R})$ denotes the set of all smooth functions f such that $\|V_1 f\|_\infty$ is finite.

We shall denote by $J_t = J_t^x = \frac{\partial}{\partial x} X_t^{(x)}$ the one dimensional process which denotes the derivative of $X_t^{(x)}$ with respect to x; this exists by [61, Theorem 7.3] and can be viewed as the solution of

$$dJ_t^x = V'_{0,\text{Itô}}(X_t^{(x)})J_t^x dt + \sqrt{2}V_1'(X_t^{(x)})J_t^x dB_t, \quad J_0^x = 1\,. \tag{4.7}$$

With this notation in place, we rewrite derivatives of the semigroup in terms of derivatives of the process $X_t^{(x)}$.

Lemma 4.1.3. *Let $\mathcal{P}_t$ be the semigroup generated by the SDE* (4.3) *and assume that the LOAC* (1.17) *is satisfied by the vector fields in* (4.3) *with a function $\lambda(x)$ such that $\lambda(x) \geq -\kappa$ for every $x \in \mathbb{R}$, for some $\kappa \in \mathbb{R}$ (note that κ need not be negative). Then*

$$V_1 \mathcal{P}_t f(x) = \mathbb{E}[f'(X_t^{(x)})J_t V_1(x)] \tag{4.8}$$

for every $x \in \mathbb{R}$ and $f \in C_{V_1}(\mathbb{R})$. For clarity we emphasize that here $f'(X_t^{(x)})$ denotes the derivative of f evaluated at $X_t^{(x)}$.

Proof. Fix $f \in C_{V_1}(\mathbb{R})$ and fix some initial condition $x \in \mathbb{R}$ then by the chain rule we have

$$V_1(f(X_t^{(x)})) = V_1(x) f'(X_t^{(x)}) J_t.$$

Now we can take expectations to obtain

$$\mathbb{E}\left[V_1(f(X_t^{(x)}))\right] = \mathbb{E}\left[V_1(x)f'(X_t^{(x)})J_t\right]. \tag{4.9}$$

At the end of the proof of Theorem 4.1.4 we justify swapping the expectation and the derivative on the left hand side of the above equality. After doing so we have (4.8). □

Let us introduce the two parameter random process $\{\Gamma_{s,t}\}_{0\leq s\leq t}$, defined as follows:

$$\Gamma_{s,t} = \left|f'(X_t^{(x)})J_tJ_s^{-1}V_1(X_s^{(x)})\right|^2.$$

The significance of the process $\Gamma_{s,t}$ will be more clear in view of (4.11). For the time being notice that by (4.8) we have

$$|V_1\mathcal{P}_tf(x)|^2 \leq \mathbb{E}\left[\left|f'(X_t^{(x)})J_tV_1(x)\right|^2\right] = \mathbb{E}[\Gamma_{0,t}],$$

and moreover, (using that f belongs to $C_{V_1}(\mathbb{R})$) we may estimate $\Gamma_{t,t}$ by

$$\Gamma_{t,t} = |V_1f(X_t^{(x)})|^2 \leq \|V_1f\|_\infty^2.$$

Hence to prove (4.6) it is sufficient to prove the following inequality

$$\mathbb{E}[\Gamma_{0,t}] \geq \mathbb{E}\left[\exp\left(-2\int_0^t \lambda(X_s^{(x)})ds\right)\Gamma_{t,t}\right]. \tag{4.10}$$

Before proving (4.10), we shall introduce some more notation. For each $\omega \in \Omega$, $s \leq t$ we may define the random flow map $\Phi_{s,t} : \mathbb{R} \to \mathbb{R}$ by

$$\Phi_{s,t}(x) := X_t^{(s,x)}, \quad t \geq s \geq 0.$$

Here $X_t^{(s,x)}$ denotes the solution to (1.1) given that $X_s^{(s,x)} = x$. It is shown in [61] that for almost all $\omega \in \Omega$, $\Phi_{s,t}$ is a well-defined diffeomorphism from $\mathbb{R}$ to $\mathbb{R}$ and we shall denote by $J_{s,t}$ the derivative $\Phi'_{s,t}(X_s^{(x)})$. By differentiating the identity

$X_t^{(x)} = \Phi_{s,t}(X_s^{(x)})$ with respect to x, we have $J_t = \Phi'_{s,t}(X_s^{(x)})J_s$ and hence

$$J_{s,t} = \Phi'_{s,t}(X_s^{(x)}) = J_t J_s^{-1}.$$

Analogously, if $f_{s,t}(\cdot) := f(\Phi_{s,t}(\cdot))$, then $f_{s,t}(X_s^{(x)}) := f(\Phi_{s,t}(X_s^{(x)}))$, so that $f'_{s,t}(X_s^{(x)}) = f'(X_t^{(x)})J_t J_s^{-1}$ and we may write

$$\Gamma_{s,t} = \left|V_1 f_{s,t}(X_s^{(x)})\right|^2 . \tag{4.11}$$

Theorem 4.1.4. *Let $X_t^{(x)}$ be the solution to SDE* (4.3) *and suppose that the Local Obtuse Angle Condition* (1.17) *is satisfied by the vector fields appearing in* (4.3) *with $\lambda(x) \geq -\kappa$ for every $x \in \mathbb{R}$ and some $\kappa \in \mathbb{R}$. Then* (4.6) *holds.*

Note 4.1.5. Some clarifications on the statement of the above theorem.

- Because the initial profile $f(x)$ is assumed to be smooth and the coefficients of the equation are smooth as well, the derivative $V_1\mathcal{P}_t f$ always makes sense. Corollary 4.1.6 below deals with the case in which f is not smooth but just continuous and bounded.

- As we have already explained, we will require further conditions to ensure that the right hand side of (4.6) decays exponentially. We will give conditions under which the right hand side of (4.6) decays exponentially in Section 4.2.

- In principle similar arguments to those used in the proof of Theorem 4.1.4 can be used to obtain estimates on higher order derivatives. In this thesis, we will demonstrate this idea for a class of examples in Lemma 4.1.10 below.

□

Proof of Theorem 4.1.4. We will use [35, Equation (2.63)] which, in our notation and setting can be written as

$$d(J_t^{-1}V(X_t^{(x)})) = J_t^{-1}[V_0, V](X_t^{(x)})dt + \sqrt{2}J_t^{-1}[V_1, V](X_t^{x}) \circ dB_t, \tag{4.12}$$

where $V : \mathbb{R} \to \mathbb{R}$ is any smooth vector field. By taking $V = V_1$ in (4.12), we obtain

$$d\left(J_t^{-1}V_1(X_t^{(x)})\right)^2 = 2J_t^{-1}[V_0, V_1](X_t^{(x)})V_1(X_t^{(x)})(J_t^{-1})dt.$$

Integrating from 0 to s and multiplying both sides by $f'(X_t^{(x)})^2 J_t^2$ one gets

$$\begin{aligned}\left|f'(X_t^{(x)})J_tJ_s^{-1}V_1(X_t^{(x)})\right|^2 &= \left|f'(X_t^{(x)})J_tV_1(x)\right|^2 \\ &+ 2\int_0^s f'(X_t^{(x)})J_tJ_r^{-1}[V_0, V_1](X_r^{(x)})V_1(X_r^{(x)})J_r^{-1}J_tf'(X_t^{(x)})dr.\end{aligned}$$

Now we may apply (1.17) and obtain

$$\begin{aligned}\left|f'(X_t^{(x)})J_tJ_s^{-1}V_1(X_t^{(x)})\right|^2 &\geq \left|f'(X_t^{(x)})J_tV_1(x)\right|^2 \\ &+ 2\int_0^s \lambda(X_r^{(x)})\left|f'(X_t^{(x)})J_tJ_r^{-1}V_1(X_r^{(x)})\right| dr.\end{aligned}$$

We can rewrite this in terms of $\Gamma_{s,t}$ as

$$\Gamma_{s,t} \geq \Gamma_{0,t} + 2\int_0^s \lambda(X_r^{(x)})\Gamma_{r,t}dr.$$

From this we obtain

$$\frac{\partial}{\partial s}\left(\exp\left(-2\int_0^s \lambda(X_r^{(x)})dr\right)\Gamma_{s,t}\right) \geq 0.$$

That is,

$$\exp\left(-2\int_0^s (\lambda(X_r^{(x)}))dr\right)\Gamma_{s,t} \geq \Gamma_{0,t}. \tag{4.13}$$

Taking expectations and setting $s = t$ one obtains (4.10).

It remains to justify that we may swap the expectation and the derivative on the left hand side of (4.9). This follows from the dominated convergence theorem provided we have

$$\sup_{x\in\mathbb{R}}|V_1(f(X_t^{(x)}))|$$

is bounded by a constant which may depend on t. By setting $s = t$ in (4.13) we have

$$|V_1(f(X_t^{(x)}))|^2 = \Gamma_{0,t} \le \exp\left(-2\int_0^t (\lambda(X_r^{(x)}))dr\right)\Gamma_{t,t}.$$

We may bound the right hand side using $-\lambda(x) \le \kappa$ and $\Gamma_{0,t} \le \|V_1 f\|^2$, this gives

$$|V_1(f(X_t^{(x)}))|^2 = \Gamma_{0,t} \le e^{2\kappa t}\|V_1 f\|^2.$$

This concludes the proof. □

We now state a simple consequence of Theorem 4.1.4, Corollary 4.1.6. We then give some simple examples to which Theorem 4.1.4 can be applied. Before stating Corollary 4.1.6 we observe that (4.6) holds for smooth functions only. Corollary 4.1.6 allows one to state an analogous result for functions f which are only continuous and bounded. We start by recalling a well-known short-time smoothing result: for any compact set K there is a constant $c = c(K)$ such that

$$|V_1\mathcal{P}_t f(x)| \le \frac{c(K)}{t}\|f\|_\infty, \quad f \in C_b(\mathbb{R}), t \in (0,1). \tag{4.14}$$

Using the above and the semigroup property, by the same argument as in [32, Note 3.2], we obtain what follows. Such smoothing estimates hold under very general assumptions on the coefficients of the SDE, for example they do hold under the UFG condition, see Appendix A.1.4 for an account on the matter (note that UFG processes include both elliptic and uniformly hypoelliptic processes).

Corollary 4.1.6. *Consider the SDE* (4.3) *and assume that the LOAC* (1.17) *and the smoothing property* (4.14) *hold. Then, for any $t_0 > 0$ and compact set K we can find a constant $c_{t_0,K}$ such that*

$$|V_1\mathcal{P}_t f(x)| \le c_{t_0,K}\mathbb{E}\left[\exp\left(-2\int_0^{t-t_0}\lambda(X_r^{(x)})dr\right)\right]^{\frac{1}{2}}\|f\|_\infty, \quad \forall x \in K, f \in C_b(\mathbb{R}^N), t \ge t_0.$$

The examples below illustrate the situation in which the OAC (1.16) does not hold but the LOAC (1.17) does.

Example 4.1.7. Consider the SDE

$$dX_t = -\arctan(X_t)dt + \sqrt{2}dW_t. \tag{4.15}$$

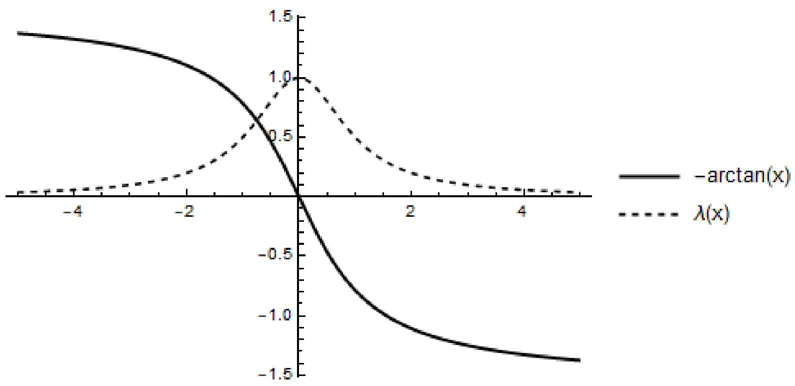


Figure 4.1: A plot of $V_0(x)$ and $\lambda(x)$ for the SDE (4.15).

In this case $N = d = 1$ and we have $V_0(x) = -\arctan(x)$, $V_1(x) = 1$. Then the LOAC (1.17) is satisfied with

$$\lambda(x) = -\frac{[V_1, V_0](x)V_1(x)}{V_1(x)^2} = \frac{1}{1+x^2}.$$

In Figure 4.1 is a plot of V_0 and λ. Notice that because $\lambda(x)$ converges to 0 as x tends to $\pm\infty$ the Obtuse Angle Condition (1.16) does not hold. By Theorem 4.1.4 we have

$$|\partial_x \mathcal{P}_t f(x)| \leq \mathbb{E}\left[\exp\left(-2\int_0^t \frac{1}{1+(X_r^{(x)})^2} dr\right)\right]^{\frac{1}{2}} \|\partial_x f\|_\infty. \tag{4.16}$$

We will continue investigating this SDE in Example 4.3.3 where we will show that right hand side of (4.16) decays exponentially. $\square$

Example 4.1.8. Consider the one-dimensional SDE

$$dX_t = -\sin(X_t)dt + \sqrt{2}\cos(X_t) \circ dB_t. \tag{4.17}$$

In this case we have $V_0(x) = -\sin(x)\partial_x$, $V_1(x) = \cos(x)\partial_x$, so that $[V_1, V_0] = -\partial_x$ and the LOAC (1.17) is satisfied with

$$\lambda(x) = \frac{1}{\cos(x)}.$$

Here (2.8) is not satisfied, indeed λ is negative for $x \in (\pi/2, 3\pi/2)$ and not defined for $x = k\pi + \pi/2$ for any $k \in \mathbb{Z}$. We also have that $\lambda(x) \geq 1$ for $x \in (-\pi/2, \pi/2)$. On the other hand, if $x \in (-\pi/2, \pi/2)$ then $X_t^{(x)} \in (-\pi/2, \pi/2)$, this can be seen directly from the SDE (4.17) or see Excursus 3.1.5. Therefore by Theorem 4.1.4 we have

$$|V_1\mathcal{P}_t f(x)| \le e^{-t}\|V_1 f\|_\infty, \quad \forall x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Note that although we stated Theorem 4.1.4 for λ finite, the proof still holds in this case since $\mathbb{P}(X_t^{(x)} \notin (-\pi/2, \pi/2)) = 0$ for any $x \in (-\pi/2, \pi/2)$. □

Note 4.1.9. To simplify the discussion, in this note we consider again the simple setting in which $d = 1$ in (1.1) as what we want to explain is independent of the number of Brownian motions driving the dynamics. In [32] a Bakry-Emery type technique is used to prove that the OAC (1.16) implies estimates of the form (2.9). The argument used there (and in related literature) is a Gronwall-type argument and it fails if $\lambda = \lambda(x)$, i.e. if (1.17) holds in place of (1.16). To explain why this is the case, we briefly recap the backbone of the argument used in [32] (and in related literature, see e.g. [25, 46, 49]): let

$$\Gamma(f) := |V_1 f(x)|^2.$$

(Note that the above function $\Gamma(f)$ is the analogous of our $\Gamma_{s,t}$ in Theorem 4.1.4). The aim is to show the following inequality:

$$\partial_s \mathcal{P}_{t-s}\Gamma(\mathcal{P}_s f(x)) \le -\lambda \mathcal{P}_{t-s}\Gamma(\mathcal{P}_s f(x)). \tag{4.18}$$

Indeed, if the above holds, then the Gronwall lemma gives

$$\mathcal{P}_{t-s}\Gamma(\mathcal{P}_s f(x)) \le e^{-\lambda}\mathcal{P}_t\Gamma(f(x))$$

and the desired exponential decay of the derivative of the semigroup in the direction V_1 is obtained by just calculating the above in $s = t$. In order to obtain (4.18) it is sufficient to prove (see [32]) the following inequality

$$(\partial_t - \mathcal{L})\Gamma(\mathcal{P}_t f(x)) \le -\lambda\Gamma(\mathcal{P}_t f(x)).$$

To prove the above the OAC was employed. In the case when $\lambda = \lambda(x)$ we can follow the same argument and this time we obtain

$$(\partial_t - \mathcal{L})\Gamma(\mathcal{P}_t f(x)) \leq -\lambda(x)\Gamma(\mathcal{P}_t f(x)).$$

However instead of (4.18) this implies

$$\partial_s \mathcal{P}_{t-s}\Gamma(\mathcal{P}_t f) \leq -\mathcal{P}_{t-s}(\lambda(x)\Gamma(\mathcal{P}_s f(x))).$$

Clearly, if $\lambda(x)$ is uniformly bounded below, then one can go on and use the previous argument again. If this is not the case then the Gronwall argument is no longer applicable. $\square$

Lemma 4.1.10. *Consider the one-dimensional SDE*

$$dX_t = b(X_t)dt + \sqrt{2}dB_t. \tag{4.19}$$

With the notation set so far, for this example we have $V_0 = U_0 = b(x)\partial_x$ *and* $V_1 = \partial_x$. *Here* $b : \mathbb{R} \to \mathbb{R}$ *is a smooth function with bounded derivatives of all orders (but* $b(x)$ *itself is not assumed to be bounded). Then there exist a constant* $\overline{\lambda} > 0$ *and a function* $u : \mathbb{R}^N \to \mathbb{R}$ *such that for all* $f \in C_b^4(\mathbb{R}^N)$ *we have*

$$\sum_{k=1}^{4} |\partial_x^4 \mathcal{P}_t f(x)| \leq u(x)e^{-\lambda t}\|f\|_{C_b^4}. \tag{4.20}$$

for the semigroup generated by the process (4.19) *provided the drift* $b(x)$ *has bounded second, third, and fourth order derivatives,* $b'(x) \leq 0$ *and there is a positive constant* $C > 0$ *such that*

$$\mathbb{E}\left[\exp\left(\int_0^t b'(X_s^{(x)})ds\right)\right] \leq u(x)e^{-Ct} \tag{4.21}$$

for some positive function $u : \mathbb{R} \to \mathbb{R}$.

The proof of this lemma can be found in Appendix A.2.2.

Note 4.1.11. We can in fact obtain a similar estimate for an arbitrarily high order of derivatives by the same process. The significance of having forth order derivatives decay exponentially is that it is shown in [2] that if b is bounded and satisfies (4.21) then the weak error of the Euler scheme converges to zero uniformly in time.

4.2 Estimates for functionals of the occupation measure

In Section 4.1 we gave conditions under which the estimate (4.6) holds. To obtain exponential decay of derivatives it remains to find conditions under which there exists a constant $C > 0$ and a function $u : \mathbb{R} \to \mathbb{R}$ with

$$\mathbb{E}\left[\exp\left(-2\int_0^t \lambda(X_s^{(x)})ds\right)\right] \leq u(x)e^{-2Ct}. \tag{4.22}$$

This is the scope of this section. As we mentioned in the introduction, for a matter of notational consistency we present the results of this section in the case in which $N = 1, d = 1$, but everything we say here is more general. Clearly, a case under which the estimate (4.22) follows immediately is the one in which the function λ is bounded below by a positive constant i.e. $\lambda(x) \geq \lambda_0 > 0$. In particular, our results hold in the case considered in [32].

We can consider the weaker situation in which $\lambda \geq 0$ and there is some set F on which $\lambda(x) \geq \lambda_F > 0$ for some positive constant λ_F. Then we require that the process spends a positive proportion of time in the set F, see Note 4.1.2. More precisely, the following holds.

Proposition 4.2.1. *Let $X_t^{(x)}$ be the solution of the SDE* (4.3)*. Suppose that there exists some set $F \subseteq \mathbb{R}$ and a constant $r > 0$ such that*

$$\frac{1}{t}\int_0^t \mathbb{1}_F(X_s^{(x)})ds \geq r \quad \mathbb{P}-a.s, \ \textit{for all } x \in \mathbb{R}.$$

Let $\lambda : \mathbb{R} \to \mathbb{R}$ be any function[15] *such that $\lambda(x) \geq 0$ for every $x \in \mathbb{R}$ and there is a positive constant λ_F such that $\lambda(x) \geq \lambda_F > 0$ for all $x \in F$. Then, for all $t \geq 0$, we have*

$$\mathbb{E}\left[\exp\left(-2\int_0^t \lambda(X_s^{(x)})ds\right)\right] \leq \mathbb{E}\left[\exp\left(-2\int_0^t \lambda_F\mathbb{1}_F(X_s^{(x)})ds\right)\right] \leq \exp\left(-2r\lambda_F t\right). \tag{4.23}$$

Moreover, let $\mathcal{P}_t$ be the semigroup associated with (4.3)*. If, additionally, the function λ satisfies the assumptions of Theorem 4.1.4, combining* (4.6) *and* (4.23)*, one*

[15] At this stage we do not assume that $\lambda(x)$ is the function appearing in the LOAC.

obtains

$$|V_1\mathcal{P}_t f(x)| \le e^{-r\lambda_F t}\|Vf\|_\infty, \quad \textit{for all } f \in C_{V_1}(\mathbb{R}), x \in \mathbb{R}, t \ge 0.$$

We can view this as a form of recurrence. We can revisit this idea by using the large deviation principle for occupation measures introduced by Donsker and Varadhan. In a series of papers [52]-[55] Donsker and Varadhan introduced conditions to obtain a large deviation principle (LDP) for the *occupation measure* of $X_t^{(x)}$, i.e. the random measure

$$l_t^x(\omega, A) = \frac{1}{t}\int_0^t \mathbb{1}_A(X_s^{(x)}(\omega))ds. \tag{4.24}$$

We briefly recall that the occupation measure l_t^x satisfies a large deviation principle if there exists a rate function $I : \mathcal{M} \to \mathbb{R}$ such that

$$\limsup_{t\to\infty} \frac{1}{t}\log(\mathbb{P}(l_t^x \in C)) \le -\inf_{\mu\in C} I(\mu), \quad \text{for all closed sets } C \subseteq \mathcal{M} \tag{4.25}$$

$$\limsup_{t\to\infty} \frac{1}{t}\log(\mathbb{P}(l_t^x \in \mathcal{O})) \le -\inf_{\mu\in \mathcal{O}} I(\mu), \quad \text{for all open sets } \mathcal{O} \subseteq \mathcal{M}. \tag{4.26}$$

Note that $(\Omega, \mathcal{F}, \mathbb{P})$ is the probability space on which the stochastic process X_t is defined. Here $\mathcal{M}$ is endowed with the weak topology. We do not give details on this notion and refer the reader to [52]-[55]. For our purpose it is important to recall that if the occupation measure satisfies a LDP with rate function $I : \mathcal{M} \to \mathbb{R}$ (here $\mathcal{M}$ denotes the space of probability measures on $\mathbb{R}$) then for any weakly continuous functional[16] $\Psi : \mathcal{M} \to \mathbb{R}$ and compact set $K \subseteq \mathbb{R}$, we have

$$\lim_{t\to\infty} \frac{1}{t}\log\sup_{x\in K}\int_\Omega [\exp(-t\Psi(l_t^x(\omega,\cdot))]\,\mathbb{P}(d\omega) = -\inf_{\mu\in\mathcal{M}}[\Psi(\mu) + I(\mu)]. \tag{4.27}$$

If $\lambda : \mathbb{R} \to \mathbb{R}$ is a continuous function we may take $\Psi : \mathcal{M} \to \mathbb{R}$ to be

$$\Psi(\mu) = \int_\mathbb{R} \lambda(y)\mu(dy).$$

[16] A functional $\Psi : \mathcal{M} \to \mathbb{R}$ is weakly continuous if given a sequence of measures μ_k which converge to a probability measure μ in the weak topology then $\Psi(\mu_k)$ converges to $\Psi(\mu)$.

Then (4.27) becomes

$$\lim_{t\to\infty}\frac{1}{t}\log\sup_{x\in K}\mathbb{E}\left[\exp\left(-\int_0^t \lambda(X_s^{(x)})ds\right)\right] = -\inf_{\mu\in\mathcal{M}}\left[\int_{\mathbb{R}}\lambda(y)\mu(dy)+I(\mu)\right]. \quad (4.28)$$

Proposition 4.2.2. *Let $X_t^{(x)}$ be the solution of the SDE* (4.3)*. Suppose the occupation measure* (4.24) *satisfies a LDP with rate function I and assume there is a continuous function $\lambda:\mathbb{R}\to\mathbb{R}$ which satisfies* (1.17)*. If*

$$\inf_{\mu\in M}\left[\int 2\lambda d\mu + I(\mu)\right] > 0 \quad (4.29)$$

then for each compact set $K\subset\mathbb{R}$ there exists a constant $C_K>0$ such that

$$\sup_{x\in K}|V_1\mathcal{P}_t f(x)| \le C_K e^{-\lambda_0 t}\|V_1 f\|_\infty, \quad \forall f\in C_{V_1}(\mathbb{R}), \quad (4.30)$$

for some $\lambda_0>0$ (independent of the compact set K). We recall that in [55] a set of conditions is given in order for the occupation measure to satisfy a LDP. These are stated in Hypothesis 4.2.3 below.[17]

Hypothesis 4.2.3. Let $X_t^{(x)}$ be the solution of the SDE (4.3) and $\mathcal{L}$ be the corresponding generator.

1. There exists a function $\Xi:\mathbb{R}\to\mathbb{R}$ and a sequence $u_n\in D(\mathcal{L})$ (here $D(\mathcal{L})$ denotes the domain of the operator $\mathcal{L}: D(\mathcal{L})\subseteq C_b(\mathbb{R})\to C_b(\mathbb{R})$) such that the following properties hold:

   (1a) The set $\{x\in\mathbb{R}:\Xi(x)\ge \ell\}$ is compact for each $\ell\in\mathbb{R}$;

   (1b) For all $n\in\mathbb{N}, x\in\mathbb{R}$ we have $u_n\ge 1$;

   (1c) For each compact set $W\subseteq\mathbb{R}$,

   $$\sup_{x\in W}\sup_{n\in\mathbb{N}} u_n(x)<\infty;$$

   (1d) For each $x\in\mathbb{R}$,

   $$\lim_{n\to\infty}\frac{\mathcal{L}u_n(x)}{u_n(x)} = \Xi(x); \quad (4.31)$$

[17] Here such conditions are stated in our notation and (one-dimensional) setting.

(1e) For some $A < \infty$

$$\sup_{n \in \mathbb{N}, x \in \mathbb{R}^N} \frac{\mathcal{L}u_n(x)}{u_n(x)} \leq A; \tag{4.32}$$

2. Assume that the law of $X_t^{(x)}$ admits a density $p(t, x, y)$ with respect to Lebesgue measure on $\mathbb{R}$ such that for all $x \in \mathbb{R}$:

   (2a) $p(1, x, y) > 0$ for almost all $y \in \mathbb{R}$

   (2b) The map $x \mapsto p(1, x, \cdot)$ is a continuous map from $\mathbb{R}$ to L^1.

Note 4.2.4. Let us comment on the above hypothesis.

- The first set of assumptions, Hypothesis 4.2.3 (1a)–(1e), are sufficient for an upper bound in the large deviation principle to hold, i.e. there is a rate function $I : \mathcal{M} \to \mathbb{R}$ such that (4.25) holds. One strategy to construct the sequence u_n appearing in Hypothesis 4.2.3 is as follows: first we find a pair of functions $u, \Xi : \mathbb{R} \to \mathbb{R}$ such that

  $$\mathcal{L}u(x) = \Xi(x)u(x)$$

  and we require that $u \geq 1$, $\Xi(x)$ is bounded above but tends to $-\infty$ as $|x| \to \infty$; we then construct the sequence $\{u_n\}_{n \in \mathbb{N}}$ by defining $u_n(x) = u(n\theta(x/n))$ where θ is a smooth function such that $\theta(-x) = -\theta(x)$ and

  $$\theta(y) = \begin{cases} y, & 0 \leq y \leq 1; \\ \text{smooth and increasing}, & 1 \leq y \leq 2; \\ 2, & y \geq 2. \end{cases}$$

  The second set of assumptions, Hypothesis 4.2.3 (2), are sufficient for a lower bound in the large deviation principle, i.e. under Hypothesis 4.2.3 (2a)–(2b) there is a rate function $I : \mathcal{M} \to \mathbb{R}$ such that (4.26) holds. Note that in the case when (4.3) satisfies a uniform ellipticity condition, i.e. there is some constant $\nu > 0$ such that $V_1(x) \geq \nu > 0$ for all $x \in \mathbb{R}$, then Hypothesis 4.2.3 (2a)–(2b) are satisfied (in contrast, under the weaker UFG condition this later set of assumptions is not satisfied).

- Note that Hypothesis 4.2.3 (1a) implies that Ξ is not bounded below, while

Hypothesis 4.2.3 (1d) and Hypothesis 4.2.3 (1e) imply that Ξ is bounded above by A.

$\square$

By [54, Theorem 7.2 and Theorem 8.1] under Hypothesis 4.2.3 the limit in (4.28) holds with

$$I(\mu) = \sup_{u \in D(\mathcal{L}), u > 0} -\int_{\mathbb{R}} \frac{\mathcal{L}u}{u} d\mu. \tag{4.33}$$

In order to prove that (4.6) implies (2.9) when Hypothesis 4.2.3 is satisfied it remains to show that the right hand side of (4.28) is positive. Note that by Fatou's lemma and (4.32) we have

$$I(\mu) \geq \liminf_{n \to \infty} \int_{\mathbb{R}} -\frac{\mathcal{L}u_n}{u_n} d\mu \geq \int_{\mathbb{R}} -\Xi d\mu.$$

In particular

$$\inf_{\mu \in M} \left[\int 2\lambda d\mu + I(\mu) \right] \geq \inf_{\mu \in M} \int_{\mathbb{R}} (2\lambda - \Xi) d\mu.$$

We have therefore proven the following.

Proposition 4.2.5. *Let $X_t^{(x)}$ be the solution of the SDE* (4.3) *with $X_0^{(x)} = x$. Assume that Hypothesis 4.2.3 holds and there exists some continuous function $\lambda : \mathbb{R} \to \mathbb{R}$ satisfying* (1.17) *and a constant $\lambda_0 > 0$ such that $2\lambda(x) - \Xi(x) \geq 2\lambda_0$ for all $x \in \mathbb{R}$. Then for each compact set $K \subseteq \mathbb{R}$ there is a constant C_K such that*

$$\sup_{x \in K} \mathbb{E} \left[\exp \left(-2 \int_0^t \lambda(X_r^{(x)}) dr \right) \right] \leq C_K e^{-2\lambda_0 t}, \quad \forall t \geq 0.$$

Note 4.2.6. Note that since Ξ tends to $-\infty$ as $x \to \pm\infty$, for $|x|$ sufficiently large we have $\Xi(x) < 0$ in which case the condition $2\lambda - \Xi \geq 2\lambda_0$ is weaker than the requirement that $\lambda \geq \lambda_0 > 0$ for $|x|$ sufficiently large. In Example 4.3.2 we illustrate a case in which we are able to find a constant $\lambda_0 > 0$ such that $2\lambda(x) - \Xi(x) > 2\lambda_0$ for all $x \in \mathbb{R}$ but $\lambda(x_0) < 0$ for some $x_0 \in \mathbb{R}$.

Hypothesis 4.2.3 is stronger than we require in order to control $|V_1 \mathcal{P}_t f(x)|$. Indeed all we require is an upper bound for the left hand side of (4.28) and we can achieve this under the following conditions.

Hypothesis 4.2.7. With the same notation and setting as Hypothesis 4.2.3, there exist a function $\Xi : \mathbb{R} \to \mathbb{R}$ and a sequence $u_n \in D(\mathcal{L})$ such that conditions (1b) - (1e) of Hypothesis 4.2.3 hold.

In particular we are no longer assuming that Ξ is unbounded from below which was required by Hypothesis 4.2.3 (1a) (see Note 4.2.4), instead we require the existence of some constant $\lambda_0 > 0$ such that

$$2\lambda(x) - \Xi(x) \geq 2\lambda_0 > 0. \tag{4.34}$$

Hypothesis 4.2.7 is weaker than Hypothesis 4.2.3 and the price we pay is that (4.34) is harder to satisfy than when Ξ was unbounded, however we will see in Example 4.3.3 that Hypothesis 4.2.7 is satisfied although Hypothesis 4.2.3 is not.

Theorem 4.2.8. *Assume that Hypothesis 4.2.7 holds for the SDE* (4.3) *and suppose there exists a continuous function $\lambda : \mathbb{R} \to \mathbb{R}$ such that* (4.34) *holds (the function Ξ appearing in* (4.34) *is as in Hypothesis 4.2.7). Then* (4.22) *holds with $u(x) := \liminf_{n\to\infty} u_n(x)$ where $\{u_n\}$ is the sequence appearing in Hypothesis 4.2.7.*

Moreover, let $\mathcal{P}_t$ be the semigroup associated with (4.3)*. If, additionally, the function λ satisfies the assumptions of Theorem 4.1.4, combining* (4.6) *and* (4.22)*, one obtains*

$$|V_1\mathcal{P}_t f(x)|^2 \leq u(x)e^{-2\lambda_0 t}\|V_1 f\|_\infty^2, \quad \textit{for all } f \in C_{V_1}(\mathbb{R}), x \in \mathbb{R}, t \geq 0\,, \tag{4.35}$$

where λ_0 is as in (4.34)*.*

Proof of Theorem 4.2.8. Define

$$\psi_n(x,t) = \mathbb{E}\left[u_n(X_t^{(x)})\exp\left(-\int_0^t \frac{\mathcal{L}u_n(X_s^{(x)})}{u_n(X_s^{(x)})}ds\right)\right].$$

By the Feynmann Kac formula, ψ_n solves the initial value problem

$$\begin{cases} \frac{\partial\psi_n}{\partial t} = \mathcal{L}\psi_n - \frac{\mathcal{L}u_n}{u_n}\psi_n \\ \psi_n(x,0) = u_n(x). \end{cases} \tag{4.36}$$

Note that u_n is also a stationary solution to this PDE, indeed

$$\mathcal{L}u_n - \frac{\mathcal{L}u_n}{u_n}u_n = \mathcal{L}u_n - \mathcal{L}u_n = 0.$$

By [60, Theorem 5.7.6] there is at most one solution to (4.36) in the class $C^{1,2}(\mathbb{R} \times [0,T];\mathbb{R})$ for each $T > 0$ and hence we have $\psi_n(x,t) = u_n(x)$, that is

$$u_n(x) = \mathbb{E}\left[u_n(X_t^{(x)}) \exp\left(-\int_0^t \frac{\mathcal{L}u_n(X_s^{(x)})}{u_n(X_s^{(x)})} ds\right)\right].$$

Using that $u_n \geq 1$ we have

$$u_n(x) \geq \mathbb{E}\left[\exp\left(-\int_0^t \frac{\mathcal{L}u_n(X_s^{(x)})}{u_n(X_s^{(x)})} ds\right)\right].$$

By Fatou's lemma

$$u(x) = \liminf_{n\to\infty} u_n(x) \geq \mathbb{E}\left[\liminf_{n\to\infty} \exp\left(-\int_0^t \frac{\mathcal{L}u_n(X_s^{(x)})}{u_n(X_s^{(x)})} ds\right)\right]$$

Now using the continuity of the function exp we can exchange the lim inf and exp

$$u(x) \geq \mathbb{E}\left[\exp\left(-\limsup_{n\to\infty} \int_0^t \frac{\mathcal{L}u_n(X_s^{(x)})}{u_n(X_s^{(x)})} ds\right)\right]$$

Again by reverse Fatou's lemma which is justified by (4.32)

$$\begin{aligned} u(x) &\geq \mathbb{E}\left[\exp\left(-\int_0^t \limsup_{n\to\infty} \frac{\mathcal{L}u_n(X_s^{(x)})}{u_n(X_s^{(x)})} ds\right)\right] \\ &= \mathbb{E}\left[\exp\left(-\int_0^t \Xi(X_s^{(x)}) ds\right)\right] \end{aligned}$$

here we have used (4.31) to justify the last line. Now using (4.34) we have

$$u(x) \geq \mathbb{E}\left[\exp\left(-\int_0^t 2\lambda(X_s^{(x)}) ds + 2\lambda_0 t\right)\right].$$

That is,

$$\mathbb{E}\left[\exp(-2\int_0^t \lambda(X_s^{(x)}) ds)\right] \leq u(x) e^{-2\lambda_0 t}$$

as required. □

4.3 Examples

Example 4.3.1. Consider again the SDE (4.19). If $b'(x) \leq -\lambda_0 < 0$ for some constant $\lambda_0 > 0$ then one can deduce exponential decay of the derivatives of the semigroup from the results of [32]. Here we prove that the derivative estimates (4.30) hold also when $b' \leq 0$. More precisely, assuming $b(x)$ is unbounded (both above and below), we show below the two following facts: i) if $b'(x) < 0$ for every x then (4.30) holds; ii) if $b'(x) \leq 0$, then the same conclusion holds, provided Hypothesis 4.2.3 is satisfied with some Ξ such that $\Xi(x) < 0$ for all x where $b'(x) = 0$. An example of a function $b(x)$ which falls in the case i) is $b(x) = \arctan(x)\log(2 + x^2)$.

For equation (4.19) we have $V_0(x) = b(x)\partial_x, V_1(x) = \partial_x$. The Local Obtuse Angle Condition (1.17) is satisfied with $\lambda(x) = b'(x)$, therefore by Theorem 4.1.4 (4.6) holds. However since b' is not necessarily uniformly bounded away from zero we do not immediately obtain (2.9); in order to obtain exponential decay we instead use the strategy of Section 4.2. In Lemma A.2.5 we show that Hypothesis 4.2.3 holds for (4.3) when $b'(x) < 0$ for all $x \in \mathbb{R}$. By using Proposition 4.2.2, in order to obtain (2.9) it is then sufficient to show

$$\lambda_0 := \inf_{\mu \in \mathcal{M}} \left[I(\mu) - \int_{\mathbb{R}} 2b' d\mu \right] > 0,$$

where we recall that I was given by (4.33). To prove the above suppose, for a contradiction, that $\lambda_0 = 0$; then there exists some sequence of probability measures $\{\mu_k\}_{k \in \mathbb{N}}$ such that

$$I(\mu_k) - \int_{\mathbb{R}} 2b'(y)\mu_k(dy) \leq \frac{1}{k}$$

for every $k \in \mathbb{N}$. Now by Markov's inequality,

$$\mu_k(\{x \in \mathbb{R} : A - \Xi(x) > A + \ell\}) \leq \frac{A - \int_{\mathbb{R}} \Xi d\mu_k}{A + \ell}$$

where Ξ and A are as in Hypothesis 4.2.3, so that $\Xi(x) \leq A$ for all $x \in \mathbb{R}$ and

Markov's inequality is applicable. By the definition of I and Fatou's lemma we have

$$I(\mu_k) \geq \liminf_{n\to\infty} \int_{\mathbb{R}} -\frac{\mathcal{L}u_n}{u_n} d\mu_k \geq \int_{\mathbb{R}} -\Xi d\mu_k.$$

This gives

$$\mu_k\{x \in \mathbb{R} : \Xi(x) \leq -\ell\} \leq \frac{A + I(\mu_k)}{A + \ell},$$

which implies that $\{\mu_k\}$ is tight since $\{x \in \mathbb{R} : \Xi(x) \leq -\ell\}$ is compact for all ℓ. By Prokhorov's theorem we may take a weakly convergent subsequence; let μ denote the limit of such a subsequence. Then

$$\int b'(y)\mu(dy) = 0. \tag{4.37}$$

However $b' < 0$ so we have a contradiction. This proves that (4.30) holds for the SDE (4.19).

By following the same reasoning as in the above, we can also consider the case when $b' \leq 0$, provided Hypothesis 4.2.3 holds for some Ξ such that $\Xi(x) < 0$ for all x where $b'(x) = 0$. Indeed by (4.37) we must have that $\mu(\{x : b' < 0\}) = 0$. Therefore if $\Xi(x) < 0$ whenever $b' = 0$ then we have

$$0 = I(\mu) - 2\mu(b') \geq \int_{\mathbb{R}} (-2b'(y) - \Xi(y))d\mu(y) = \int_{b'=0} -\Xi d\mu > 0$$

which gives again a contradiction. □

The above example gives us a class of SDEs for which (4.30) holds. We now consider a specific example in which the Local Obtuse Angle Condition (1.17) is satisfied however $\lambda(5) < 0$ so (1.16) is not satisfied.

Example 4.3.2. Consider the SDE

$$dX_t = (2\arctan(X_t - 5) - X_t)dt + \sqrt{2}dB_t. \tag{4.38}$$

For this example we will show that (4.30) holds. Indeed we have $V_0 = (2\arctan(x - 5) - x)\partial_x$, $V_1 = \partial_x$, and then (1.17) is satisfied with

$$\lambda(x) = 1 - \frac{2}{1 + (x-5)^2}.$$

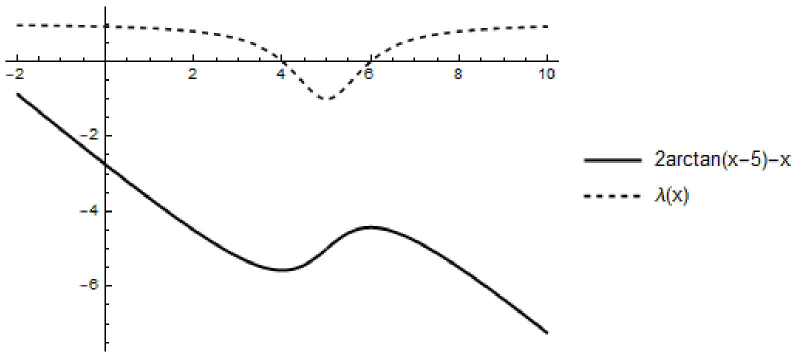


Figure 4.2: A plot of $V_0(x)$ and $\lambda(x)$ for the SDE (4.38).

Now we may apply Theorem 4.1.4 and see that (4.30) holds provided (4.22) does too. To show (4.22) we shall use Theorem 4.2.8. Note that Hypothesis 4.2.7 is satisfied by Lemma A.2.5.

Note that the function λ in this case is bounded below by -1 but does take negative values. In Figure 4.2 we plot both $V_0(x)$ and $\lambda(x)$. By Lemma A.2.5 we have that Hypothesis 4.2.7 is satisfied with $\Xi = 0.25 + 0.5(2\arctan(x-5) - x)\tanh(0.5x)$. Then by Theorem 4.2.8 we have that (4.30) follows provided we can find a $\lambda_0 > 0$ with

$$2\lambda(x) - \Xi(x) > 2\lambda_0, \qquad \text{for all } x \in \mathbb{R}.$$

In Figure 4.3 we can see there is a constant $\lambda_0 > 0$ such that $2\lambda(x) - \Xi(x) \geq 2\lambda_0$ for all $x \in \mathbb{R}$, hence by Theorem 4.1.4 and Theorem 4.2.8 we have

$$|\partial_x \mathcal{P}_t f(x)| \leq \cosh(0.5x) e^{-\lambda_0 t} \|\partial_x f\|_\infty.$$

Note that here we have used that we may take $u(x) = \cosh(0.5x)$ which follows from the proof of Lemma A.2.5 with $\alpha = 0.5$. □

Example 4.3.3. Here we continue Example 4.1.7, i.e. we consider again the SDE (4.15). Notice that (4.21) is just (4.22) with $\lambda(x) = -b'(x)$, $b(x)$ being the drift in (4.15). Therefore, to obtain (4.21), we use Theorem 4.2.8. In turn, to apply Theorem 4.2.8, we must verify that Hypothesis 4.2.7 holds. This is done in Lemma A.2.5, where we show that (4.15) satisfies Hypothesis 4.2.7 with

$$\Xi(x) = \frac{1}{4} - \frac{1}{2}\arctan(x)\tanh\left(\frac{x}{2}\right).$$

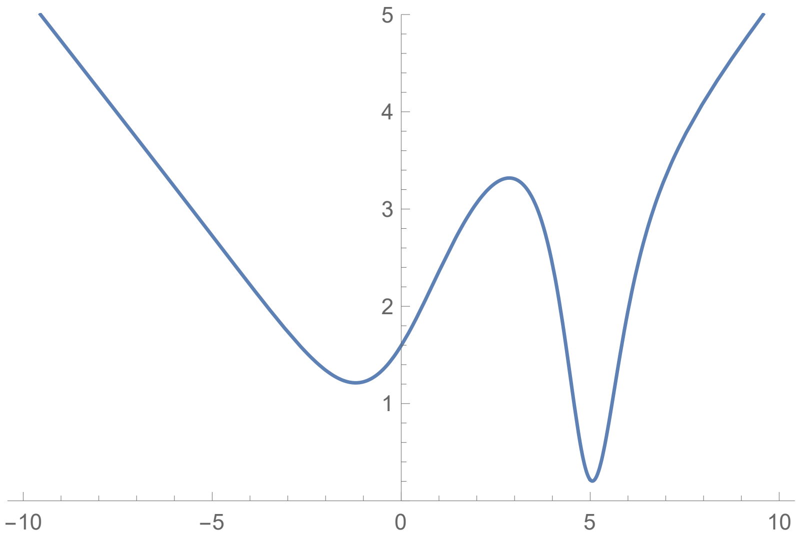


Figure 4.3: A plot of $2\lambda(x) - \Xi(x)$ for the SDE (4.38).

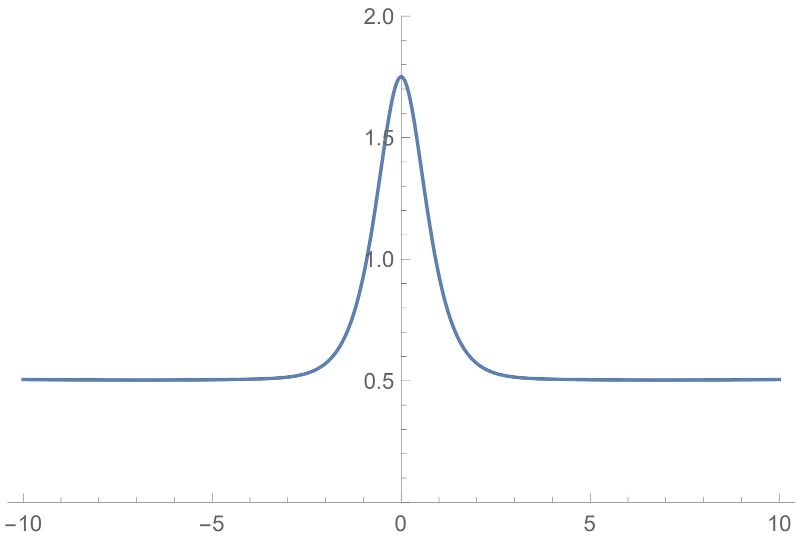


Figure 4.4: A plot of $\lambda(x) - \Xi(x)$ for the SDE (4.15).

From the proof of Lemma A.2.5 one can moreover see that (4.21) holds with $u(x) = \cosh(x/2)$. In Figure 4.4 we can see there is a constant[18] $\lambda_0 > 0$ such that $2\lambda(x) - \Xi(x) \geq 2\lambda_0$ for all $x \in \mathbb{R}$, hence by Theorem 4.2.8 we have

$$\mathbb{E}\left[\exp\left(-2\int_0^t \frac{1}{1+(X_s)^2}ds\right)\right] \leq \cosh(x/2)e^{-\lambda_0 t}.$$

Note that another consequence of (4.20) is that the SDE (4.15) decays to equilibrium exponentially fast. One can check directly that (4.15) admits an invariant

[18]One can find numerically that λ_0 is about 0.267.

measure and such an invariant measure has a density with respect to the Lebesgue measure on $\mathbb{R}$ given by

$$\mu(x) = \frac{1}{Z}\sqrt{1+x^2}e^{-x\arctan(x)},$$

where Z is a normalising constant. Then for $f \in C_b^1(\mathbb{R})$ we have

$$\begin{aligned}
\left|\mathcal{P}_t f(x) - \int_{\mathbb{R}} f(y)\mu(y)dy\right| &= \left|\mathcal{P}_t f(x) - \int_{\mathbb{R}} \mathcal{P}_t f(y)\mu(y)dy\right| \\
&= \left|\int_{\mathbb{R}} (\mathcal{P}_t f(x) - \mathcal{P}_t f(y))\mu(y)dy\right| \\
&= \left|\int_{\mathbb{R}} \int_x^y \partial_z \mathcal{P}_t f(z)dz\mu(y)dy\right| \\
&\leq \|f'\|_\infty e^{-\lambda_0 t} \int_{\mathbb{R}} \int_x^y \cosh(z/2)dz\mu(y)dy \\
&\leq K(x)\|f'\|_\infty e^{-\lambda_0 t}.
\end{aligned}$$

Here

$$K(x) = \int_{\mathbb{R}} \int_x^y \cosh(z/2)dz\mu(y)dy$$

which is finite for all $x \in \mathbb{R}$. □

Chapter 5

Long-time behaviour of UFG processes

5.1 Long time behaviour of UFG processes: the case of non-autonomous hypoelliptic diffusions

In this section we set $N = n + 1$ and study stochastic dynamics in $\mathbb{R}^N = \mathbb{R}^{n+1}$ of the form

$$dZ_t = U_0(Z_t, \zeta_t)dt + \sum_{j=1}^{d} U_j(Z_t, \zeta_t) \circ dB_t^j, \tag{5.1}$$

$$d\zeta_t = W_0(\zeta_t)dt \tag{5.2}$$

$$Z_0 = z, \quad \zeta_0 = \zeta\,. \tag{5.3}$$

In other words, we consider systems for which the representation of the form "ODE+SDE" (3.10)- (3.11) is global.[19] The above system consists of an $n-$dimensional process, $Z_t \in \mathbb{R}^n$, satisfying an SDE, equation (5.1), which is coupled with a one-dimensional autonomous ODE, (5.2). As in previous sections, $U_j : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n, j \in \{0, \ldots, d\}$ and $W_0 : \mathbb{R} \to \mathbb{R}$. The evolution of Z_t depends on the evolution of ζ_t, but the ODE solution ζ_t evolves independently of the SDE. For the purposes of this thesis,

[19] We are not claiming that this representation necessarily results from the change of coordinates presented in Section 3.1.2.

we don't think of ζ_t as representing time, but rather as representing an additional space-coordinate. However notice that if $W_0 \equiv 1$ and $\zeta(0) = 0$ then $\zeta_t = t$ and we recover a standard time-inhomogeneous setting, i.e. in this case (5.1) becomes a general time-inhomogeneous SDE, namely

$$dZ_t = U_0(Z_t, t)dt + \sum_{j=1}^{d} U_j(Z_t, t) \circ dB_t^j \,. \tag{5.4}$$

Going back to the representation of the form "ODE+SDE" (5.1)-(5.2) under consideration, if we denote by X_t the process $\mathbb{R}^N \ni X_t = (Z_t, \zeta_t)$, then X_t is the solution of an autonomous SDE. The one-parameter semigroup associated to X_t is, as usual, given by

$$(\mathcal{P}_t f)(x) := \mathbb{E}[f(X_t)|X_0 = x] = \mathbb{E}\left[f(Z_t, \zeta_t)|(Z_0, \zeta_0) = (z, \zeta)\right], \quad x = (z, \zeta) \in \mathbb{R}^{n+1},$$

for any $f \in C_b(\mathbb{R}^{n+1}; \mathbb{R})$. On the other hand one could consider the two-parameter semigroup associated with the non-autonomous process Z_t alone. Indeed, if we solve the ODE for ζ_t and substitute the solution back into the SDE for Z_t, then we can simply consider equation (5.1) rather than the whole system. To be more precise, let us denote by ζ_t^ζ the solution at time t of (5.2) with initial datum $\zeta(0) = \zeta$. That is, $\zeta_t^\zeta = e^{tW_0}\zeta$. Let also $Z_t^{s,z,\zeta}$ be the solution of the following SDE:

$$Z_t = z + \int_s^t U_0(Z_u, \zeta_u^\zeta)du + \sum_{j=1}^{d} \int_s^t U_j(Z_u, \zeta_u^\zeta) \circ dW_u^j \,.$$

The two-parameter semigroup associated with the above non-autonomous SDE is given by

$$(Q_{s,t}^\zeta g)(z) := \mathbb{E}\left[g(Z_t^{s,z,\zeta})\right], \qquad z \in \mathbb{R}^n, s \le t \,,$$

We emphasize that this two-parameter semigroup depends on ζ, i.e. on the initial datum of the ODE. When we do not wish to stress this dependence we may just write $Q_{s,t}$. With this notation, one can equivalently rewrite the definion of $\mathcal{P}_t$ as

$$(\mathcal{P}_t f)(x) = \mathbb{E}\left[f(Z_t^{0,z,\zeta}, \zeta_t^\zeta)\right], \quad x = (z, \zeta) \in \mathbb{R}^{n+1}. \tag{5.5}$$

To make explicit the relation between the two-parameter semigroup $Q^{\zeta}_{s,t}$ and the one-parameter semigroup $\mathcal{P}_t$, fix $s \in \mathbb{R}$ and let $\hat{\zeta} = e^{sW_0}\zeta$. Notice that

$$Z_t^{0,z,\hat{\zeta}} = Z_{t+s}^{s,z,\zeta}$$

where the equality is intended in law. Therefore, for every $f \in C_b(\mathbb{R}^{n+1};\mathbb{R})$, and $z \in \mathbb{R}^n$, we have

$$\begin{aligned}(\mathcal{P}_t f)(z,\hat{\zeta}) &= \mathbb{E}\left[f(Z_t^{0,z,\hat{\zeta}}, \zeta_t^{\hat{\zeta}})\right] = \mathbb{E}\left[f(Z_t^{0,z,\hat{\zeta}}, \zeta_{t+s}^{\zeta})\right] \\ &= \mathbb{E}\left[f(Z_{t+s}^{s,z,\zeta}, \zeta_{t+s}^{\zeta})\right].\end{aligned}$$

Hence,

$$(\mathcal{P}_t f)(z,\hat{\zeta}) = (Q_{s,s+t} f(\cdot, \zeta_{t+s}^{\zeta}))(z)\,. \tag{5.6}$$

On the right hand side of the above we mean to say that the semigroup Q is acting on the function $f(\cdot, a)$ obtained by freezing the value of the last coordinate of the argument.

From now on, unless otherwise specified, we write Z_t for Z_t^{0,z,ζ_0}. With this set up in place, we can start commenting on the long-time behaviour. Heuristically, if the solution of the ODE (5.2) is unbounded, then one can't expect the process X_t to have an invariant measure (see Proposition 5.1.8)– though the process Z_t may still admit an invariant measure. So we restrict to the case in which the solution of the ODE is bounded. However, because (5.2) is a one-dimensional time-homogeneous ODE, if ζ_t is bounded then it can only either increase or decrease towards stable *stationary points* of the dynamics (a stationary point of the ODE (5.2) is a point $\bar{\zeta} \in \mathbb{R}$ such that $W_0(\bar{\zeta}) = 0$). We emphasise that there may be many such points. For these reasons, we work under the assumption that ζ_t admits a finite limit, i.e. we assume that the initial datum $\zeta_0 \in \mathbb{R}$ is such that there exists a point $\bar{\zeta} = \bar{\zeta}(\zeta_0) \in \mathbb{R}$ such that

$$\zeta_t^{\zeta_0} \to \bar{\zeta} = \bar{\zeta}(\zeta_0) \quad \text{as } t \to \infty. \tag{5.7}$$

As customary, the notation $\bar{\zeta} = \bar{\zeta}(\zeta_0)$ is to emphasise the fact that the limit point will depend on the initial datum (when we don't wish to stress such a dependence we just denote a stationary point of the ODE by $\bar{\zeta}$). The dynamics (5.1)-(5.2) will, in general, admit several invariant measures. As pointed out in the introduction, when this is the case, it is typically extremely difficult to determine the basin of attraction of each invariant measure. However in the setting of this section the basin of attraction of a given invariant measure will only depend on the behaviour of the ODE. (In the next section we will show that, despite the fact that the representation of the form "ODE+SDE" is only local for generic UFG processes, it is still the case that we can relate in a simple way the initial datum to the invariant measure to which the process is converging). Given an initial datum ζ_0 for (5.2), let $\bar{\zeta} = \bar{\zeta}(\zeta_0)$ be the corresponding limit point of the ODE dynamics, as in (5.7). Consider the SDE

$$d\bar{Z}_t = U_0(\bar{Z}_t, \bar{\zeta})\, dt + \sum_{j=1}^{d} U_j(\bar{Z}_t, \bar{\zeta}) \circ dB_t^j, \quad \bar{Z}_0 = \bar{z}, \tag{5.8}$$

with associated semigroup

$$(\bar{Q}_t g)(\bar{z}) := \mathbb{E} g(\bar{Z}_t | \bar{Z}_0 = \bar{z}), \quad \bar{z} \in \mathbb{R}^n, g \in C_b(\mathbb{R}^n)\,.$$

We will assume that the dynamics (5.8) is hypoelliptic, see Hypothesis **[H.1]** below for a more precise statement of assumptions. Moreover, under Hypothesis **[H.2]**, the semigroup $\bar{Q}_t$ admits a unique invariant measure, $\bar{\mu} = \bar{\mu}(\bar{\zeta}, \zeta_0)$ (see Lemma 5.1.4). We emphasise that the asymptotic behaviour of $\bar{Z}_t$ is independent of the initial datum $\bar{z}$, see Lemma 5.1.4.

In view of (5.7), it is reasonable to guess that the asymptotic behaviour of $Z_t = Z_t^{0,z,\zeta_0}$ is the same as the asymptotic behaviour of $\bar{Z}_t$ which is the solution of (5.8). This is the content of Theorem 5.1.5 below. Theorem 5.1.5 and Theorem 5.1.6 are the main results of this section; the former is concerned with the asymptotic behaviour of the semigroup $Q_{s,t}$, the latter describes the related asymptotic behaviour of the semigroup $\mathcal{P}_t$. We set first the assumptions used in the rest of this section and we comment on their significance in Note 5.1.2.

Hypothesis 5.1.1. With the notation introduced so far, we will consider the following assumptions:

[H.1] The vector fields $V_0 = (U_0, W_0), V_1 = (U_1, 0), \ldots, V_d = (U_d, 0)$ satisfy the UFG condition for some $m \geq 1$; moreover,

$$span\{\mathcal{R}_m\} = \text{span}\{V_{[\alpha]}(x) : \alpha \in \mathcal{A}_m\} \simeq \mathbb{R}^n, \quad \text{for every } x \in \mathbb{R}^{n+1}.$$

[**H.2**] Define the measures $q_t^{s,z}$ by $q_t^{s,z}(A) := Q_{s,t}\mathbb{1}_A(z)$, for any Borel measurable $A \subseteq \mathbb{R}^n$. Then we require that for each $z \in \mathbb{R}^n$ the family measures $\{q_t^{0,z} : t \geq 0\}$ on $\mathbb{R}^n$ is tight.

[**H.3**] The long time derivative estimate (2.9) is satisfied.

[**H.4**] The ODE (5.2) has at least one stationary point $\bar{\zeta}$ and the initial datum $\zeta_0 \in \mathbb{R}$ of (5.2) is such that (5.7) holds, for some limit point $\bar{\zeta} = \bar{\zeta}(\zeta_0)$.

Note 5.1.2. Some comments on the above assumptions, in particular on Hypothesis [**H.1**].

- We start by remarking on the obvious fact that if $X_t = (Z_t, \zeta_t)$, where Z_t, ζ_t are as in (5.1)-(5.2), then X_t solves an SDE of the form (1.1), with $V_0 = (U_0, W_0), V_1 = (U_1, 0), \dots, V_d = (U_d, 0)$.

- With the notation of Section 3.1 and Section 3.2, assumption [**H.1**] implies that the distribution $\hat{\Delta}(x)$ is n-dimensional for every $x \in \mathbb{R}^N$, with $n = N-1$. In the setting of this section, this is the maximum rank that the distribution $\hat{\Delta}$ can have (as $V_0 = (U_0, W_0)$ is not contained in $\mathcal{R}_m$ when $W_0 \neq 0$). In other words, for every $x \in \mathbb{R}^N$, the integral manifolds S_x of $\hat{\Delta}(x)$ are $(N-1)$-dimensional manifolds. Because of the particularly simple structure of the SDE, such manifolds are just hyperplanes: for $x = (z, \zeta)$, $S_x = S_{(z,\zeta)} = \{u \in \mathbb{R}^{n+1} : u = (z, \eta), \eta = \zeta, z \in \mathbb{R}^n\}$. In this explicit setting Proposition 3.2.3 is easy to check.

- To reconcile the present work with the framework of [38] and further elaborate on the meaning of Hypothesis [**H.1**], let us assume for the moment that $W_0 \equiv 1$ and that $\zeta(0) = 0$, so that (5.1) becomes a standard time inhomogeneous SDE of the form (5.4). In this case the vector fields $U_0, , \dots, U_d$ are $\mathbb{R}^n$-valued maps whose coefficients depend on time, i.e. $(z, t) \mapsto U_j(z, t) \in \mathbb{R}^n$. For simplicity, let also $n = 1$. Then V_0 acts both on space and time, while $V_1, \ldots, V_d$ act on the space coordinate z only. That is, $V_0 = U_0(z,t)\partial_z + \partial_t$ while $V_j = U_j(z,t)\partial_z$ for $j = 1, \ldots, d$, so that

$$[V_0, V_j] = [U_0, U_j] + (\partial_t U_j(z,t))\partial_z \quad j \in \{1, \ldots, d\}. \tag{5.9}$$

One can then rephrase Hypothesis **[H.1]** just in terms of the fields $U_0, \ldots, U_d$; from (5.9) it is then clear that Hypothesis **[H.1]** is equivalent to assuming that the Lie algebra

$$\mathrm{span}\{\mathfrak{L}_k^U(z,t) : k \geq 1\},$$

where $\mathfrak{L}_1^U(z,t) := \{U_1(z,t), \ldots, U_d(z,t)\}$ and, for $k > 1$, $\mathfrak{L}_k^U(z,t) := \{[U, U_j], U \in \mathfrak{L}_{k-1}^U, 1 \leq j \leq d\} \cup \{[U, U_0 + \partial_t], U \in \mathfrak{L}_{k-1}^U\}$, should be finitely generated and span $\mathbb{R}_z^n$, for every $(z,t) \in \mathbb{R}^n \times \mathbb{R}$.

Let us now go back to the general representation of the form "ODE+SDE" (5.1) - (5.2), without assuming $W_0 = 1$. Recall that in this context the vector fields U_j are $\mathbb{R}^n$-valued functions of $n+1$ variables; that is, we view them as maps $\mathbb{R}^n \times \mathbb{R} \ni (z, \zeta) \mapsto U_j(z,\zeta) \in \mathbb{R}^n$. Set again $n = 1$ just for simplicity (everything we write in this comment would be true anyway). Then, as differential operators, $U_0, \ldots, U_j$ only act on the variable z, while W_0 only acts on the variable ζ, i.e. we have the correspondence

$$U_j(z,\zeta) \longleftrightarrow U_j(z,\zeta)\partial_z \;\; \text{for } j \in \{0, \ldots, d\} \quad \text{and} \quad W_0(\zeta) \longleftrightarrow W_0(\zeta)\partial_\zeta \,.$$

For all $j \in \{1, \ldots, d\}$ one has

$$[V_0, V_j] = [U_0\partial_z, U_j\partial_z] + [W_0\partial_\zeta, U_j\partial_z] = [U_0\partial_z, U_j\partial_z] + W_0(\zeta)(\partial_\zeta U_j)\partial_z.$$

If we calculate the second term on the RHS of the above along a solution ζ_t of the ODE, we obtain

$$W_0(\zeta_t)(\partial_\zeta U_j(z,\zeta_t))\partial_z = \partial_t(U_j(z,\zeta_t))\partial_z.$$

This suggests that we may evaluate the vector fields along the solution of the

ODE and then think of them as functions of z and time t, rather than as functions of z and ζ, i.e. $\mathbb{R}^n_z \times \mathbb{R}_t \ni (z,t) \mapsto U_j(z,\zeta_t^{\zeta}) \in \mathbb{R}^n$, $j \in \{0,\dots,d\}$. If we do so, then Hypothesis **[H.1]** can be equivalently rephrased as follows: the Lie algebra

$$\bigcup_{k\geq 1} \operatorname{span}\{\mathfrak{L}_k^U(z,\zeta_t)\}$$

is finitely generated and spans $\mathbb{R}^n_z$ for every $z \in \mathbb{R}^n$ and along any solutions ζ_t of the ODE (5.2).[20]

- As is well known, Hypothesis **[H.2]** is implied by a Lyapunov-type condition; namely, if there exists some non-negative function $\varphi \in C^2(\mathbb{R}^n)$ with compact level sets and such that

$$\mathcal{L}_t\varphi(z) \leq C_1 - C_2\varphi(z), \qquad \text{for every } z \in \mathbb{R}^n, t \geq 0, \tag{5.10}$$

then **[H.2]** is satisfied. Here $\mathcal{L}_t$ is the operator

$$\mathcal{L}_t\psi(z) = U_0(z,\zeta_t)\cdot\nabla\psi(z) + \sum_{i=1}^d U_i(z,\zeta_t)\cdot\nabla(U_i(z,\zeta_t)\cdot\nabla\psi(z)),$$

where $\nabla = (\partial_{z_1},\dots,\partial_{z_n})$.

- The long time derivative estimate (2.9) does not imply tightness; in Example 5.1.9 we show that **[H.3]** does not imply **[H.2]**.

□

Note 5.1.3. As already pointed out, if $X_t = (Z_t,\zeta_t)$, where Z_t,ζ_t are given by a representation of the form "ODE+SDE" (5.1)-(5.2), then X_t solves an SDE of the form (1.1), with $V_0 = (U_0,W_0), V_1 = (U_1,0),\dots,V_d = (U_d,0)$. Hence $V_0^{(\perp)} = (0,\dots,0,W_0)$ (see definition (1.12)). We note in passing that in this case one has

$$\mathcal{Z}_t := e^{-tV_0^{(\perp)}}X_t = e^{-tV_0^{(\perp)}}(Z_t,\zeta_t) = e^{-tV_0^{(\perp)}}(Z_t,e^{tV_0^{(\perp)}}\zeta_0) = (Z_t,\zeta_0).$$

(This is not of much use at the moment, but it will help at the beginning of Section

[20] Given an initial datum, the solution of the ODE is unique. When we say that this should hold along any solutions, we mean along all the solutions that one can obtain by starting from different initial data.

5.2 to make a link between the setting of this section and the setting of the next). Therefore, while X_t belongs to the hyperplane $\mathcal{H}_{\zeta_t} := \{x \in \mathbb{R}^{n+1} : x = (z, \zeta_t), z \in \mathbb{R}^n\}$ for each $t \geq 0$, $\mathcal{Z}_t$ remains, for every $t \geq 0$, on the same hyperplane, namely the hyperplane $\mathcal{H}_{\zeta_0} := \{x \in \mathbb{R}^{n+1} : x = (z, \zeta_0), z \in \mathbb{R}^n\}$ (which is precisely the manifold $S_{x_0} = S_{(z_0,\zeta_0)}$, see second bullet point in Note 5.1.2) for every $t \geq 0$. □

Lemma 5.1.4. *Let Hypothesis 5.1.1 hold. Then the SDE* (5.8) *admits a unique invariant measure, $\bar{\mu}$. Moreover,*

$$(\bar{Q}_t g)(z) \to \int_{\mathbb{R}^n} g(z)\, \bar{\mu}(dz), \quad \textit{for every } z \in \mathbb{R}^n \textit{ and every } g \in C_b(\mathbb{R}^n)\,.$$

Proof of Lemma 5.1.4. This is completely standard and we omit it. See for example [46]. We just point out that the existence of the invariant measure comes from assumption **[H.2]** and the uniqueness is a consequence of Hypothesis **[H.3]** and Proposition 3.2.10 . □

Theorem 5.1.5. *Let Hypothesis 5.1.1 hold. In particular, let $\bar{\zeta} = \bar{\zeta}(\zeta_0)$ be a stationary point for the ODE* (5.2) *and $\bar{\mu}$ be the invariant measure of the process* (5.8)*. Then, for every $s \geq 0$,*

$$\lim_{t\to\infty} (Q^{\zeta_0}_{s,t} g)(z) = \int_{\mathbb{R}^n} g(z)\, \bar{\mu}(dz), \quad \textit{for every } z \in \mathbb{R}^n \textit{ and every } g \in C_b(\mathbb{R}^n)\,.$$

The proof of this theorem can be found after the statement of Theorem 5.1.6. Theorem 5.1.5 describes the asymptotic behaviour of the process Z_t. However, in this thesis we are interested in the process X_t. The long-time behaviour of the process X_t is described by Theorem 5.1.6 below, which is just a straightforward consequence of Theorem 5.1.5. In order to state Theorem 5.1.6, we clarify the following: while Z_t is a process in $\mathbb{R}^n$ with invariant measure(s) $\bar{\mu} = \bar{\mu}(\bar{\zeta}, \zeta_0)$ supported on $\mathbb{R}^n$, X_t is a process in $\mathbb{R}^{n+1}$; so, strictly speaking, any invariant measure of X_t is a probability measure on $\mathbb{R}^{n+1}$. However such a measure is supported on the n-dimensional hyperplane

$$\mathcal{H}_{\bar{\zeta}} := \{x \in \mathbb{R}^{n+1} : x = (z, \bar{\zeta}), z \in \mathbb{R}^n\}$$

and it is just a trivial extension of the measure $\bar{\mu}$. That is, let $\mu = \mu(\bar{\zeta}, \zeta_0)$ be the measure on $\mathbb{R}^{n+1}$ such that

$$\mu(A) = \bar{\mu}(A \cap \mathcal{H}_{\bar{\zeta}}) \quad \text{for every Borel set } A \subseteq \mathbb{R}^{n+1}. \tag{5.11}$$

In particular, $\mu(A) = \bar{\mu}(A)$ if $A \subseteq \mathcal{H}_{\bar{\zeta}}$ and $\mu(A) = 0$ if $A \cap \mathcal{H}_{\bar{\zeta}} = \emptyset$. Let $I_0(\bar{\zeta}) = \{\zeta_0 \in \mathbb{R} : \zeta_t^{\zeta_0} \to \bar{\zeta} \text{ as } t \to \infty\}$. Let also $\mathcal{I}_0 = \mathcal{I}_0(\bar{\zeta}) := \{x_0 \in \mathbb{R}^{n+1} : x_0 = (z_0, \zeta_0), \zeta_0 \in I_0(\bar{\zeta}), z_0 \in \mathbb{R}^n\}$.

Theorem 5.1.6. *Consider the process $X_t = (Z_t, \zeta_t) \in \mathbb{R}^{n+1}$ satisfying a representation the form of "SDE+ODE"* (5.1)-(5.2) *with initial condition* (5.3) *and associated semigroup $\mathcal{P}_t$, defined in* (5.5)*. Let Hypothesis 5.1.1 hold. In particular, according to Hypothesis 5.1.1* **[H.4]***, let $\bar{\zeta}$ be a (any) stationary point of the ODE* (5.2) *and $\bar{\mu}$ be the invariant measure of the corresponding process* (5.8)*; let also $\mu = \mu(\bar{\zeta}, \zeta_0)$ be the measure on $\mathbb{R}^{n+1}$ defined in* (5.11) *and supported on the hyperplane $\mathcal{H}_{\bar{\zeta}}$. Then, for every $x \in \mathcal{I}_0 = \mathcal{I}_0(\bar{\zeta})$, we have*

$$\lim_{t \to \infty} (\mathcal{P}_t f)(x) = \int_{\mathbb{R}^{n+1}} f(u)\, \mu(du) = \int_{\mathcal{H}_{\bar{\zeta}}} f(u)\, \mu(du),$$

for every $f \in C_b(\mathbb{R}^{n+1})$. The above result does not hold if $x \notin \mathcal{I}_0$; that is, $\mathcal{I}_0$ is the whole basin of attraction of the measure $\mu = \mu(\bar{\zeta}, \zeta_0)$.

We now introduce some definitions that will be needed for the proof of Theorem 5.1.5. A family $\{\nu_t\}_{t \geq 0}$ of probability measures on $\mathbb{R}^n$ is said to be an *evolution system of measures* for the two-parameter semigroup $Q_{s,t}$ if for all $0 \leq s \leq t$ and $g \in C_b(\mathbb{R}^n)$ we have

$$\int_{\mathbb{R}^n} Q_{s,t} g(z) \nu_s(dz) = \int_{\mathbb{R}^n} g(z) \nu_t(dz). \tag{5.12}$$

Let $Q^*_{s,t}$ denote the adjoint of $Q_{s,t}$ over the space $C_b(\mathbb{R}^n)$, that is

$$(Q^*_{s,t}\nu)(A) := \int_{\mathbb{R}^n} Q_{s,t} \mathbb{1}_A(z) \nu(dz), \quad \text{for any Borel measurable } A \subseteq \mathbb{R}^n.$$

Then we can write (5.12) as

$$Q^*_{s,t}\nu_s = \nu_t, \quad \text{for all } 0 \leq s \leq t.$$

Further background on evolution system of measures can be found in [40, 41].

Proof of Theorem 5.1.5. The proof is in three steps.

- *Step 1*: We first construct a tight evolution system of measures, $\{\nu_t\}_{t\geq 0}$, for the semigroup $Q_{s,t}$. To this end, take any point $z_0 \in \mathbb{R}^n$, define $\nu_0 = \delta_{z_0}$ and then let $\nu_t := Q^*_{0,t}\nu_0$. Now ν_t is an evolution system of measures; indeed,

$$Q^*_{s,t}\nu_s = Q^*_{s,t}Q^*_{0,s}\nu_0 = Q^*_{0,t}\nu_0 = \nu_t.$$

(A more general construction of the evolution system is given in [40, Section 5]). To see that $\{\nu_t\}_{t\geq 0}$ is tight, fix $\varepsilon > 0$; by Hypothesis **[H.2]** we may take a compact set $K_\varepsilon \subset \mathbb{R}^n$ such that $q_t^{0,z_0}(K_\varepsilon) \geq 1 - \varepsilon$. By definition of ν_t we then have

$$\nu_t(K_\varepsilon) = (Q^*_{0,t}\nu_0)(K_\epsilon) = Q_{0,t}\mathbb{1}_{K_\varepsilon}(z_0) \geq 1 - \varepsilon.$$

- *Step 2*: $Q_{s,t}g(z) - \nu_t(g)$ converges to zero as t tends to ∞ for all $s \geq 0, z \in \mathbb{R}^n, g \in C_b(\mathbb{R}^n)$. We defer the proof of this fact to Lemma A.2.6. Since $\{\nu_t\}_{t\geq 0}$ is tight, by Prokhorov's Theorem there exists a diverging sequence t_ℓ such that ν_{t_ℓ} converges weakly to some probability measure μ_0, as t_ℓ tends to ∞.
- *Step 3*: Show that $\mu_0 = \bar{\mu}$. We defer the proof of this equality to Lemma A.2.9. If $\mu_0 = \bar{\mu}$, then ν_t converges weakly to $\bar{\mu}$ and the claim of the theorem follows; indeed,

$$|Q_{s,t}g(z) - \bar{\mu}(g)| \leq |Q_{s,t}g(z) - \nu_t(g)| + |\nu_t(g) - \bar{\mu}(g)|.$$

The first term converges to zero by Step 2 and the second term vanishes in the limit since ν_t converges weakly to $\bar{\mu}$ as $t \to \infty$. $\square$

Note 5.1.7. The statements and proofs of Lemma A.2.6 and Lemma A.2.9 are the core of the proof of Theorem 5.1.5. The arguments used in the proofs of such lemmata are analogous in structure to those presented in [62, Section 6]. The main differences arise when dealing with the regularity of the semigroup, as [62] assumes uniform ellipticity. Lemma A.2.7 (needed to prove Lemma A.2.9) is the main place where we take care of the relaxed regularity assumptions. $\square$

Let p_t^x denote the measure defined by

$$p_t^x(A) = \mathcal{P}_t \mathbb{1}_A(x), \qquad \text{for all Borel sets } A \subseteq \mathbb{R}^{n+1}.$$

Proposition 5.1.8. *If $\zeta_t^\zeta \to \infty$ then the family of measures $\{p_t^{(z,\zeta)}\}_{t\geq 0}$ is not tight for any $z \in \mathbb{R}^n$ (hence, by Prokhorov's Theorem, there is no probability measure μ such that $\mathcal{P}_t f(z,\zeta) \to \mu(f)$, for all $f \in C_b(\mathbb{R}^{n+1})$).*

Proof of Proposition 5.1.8. Fix $z \in \mathbb{R}^n$ and let $x = (z,\zeta) \in \mathbb{R}^{n+1}$. Assume by contradiction that $\{p_t^x\}_{t\geq 0}$ is tight. Then, for any fixed $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset \mathbb{R}^{n+1}$ such that $p_t^x(K_\varepsilon) > 1-\varepsilon$ for all $t \geq 0$. Since K_ε is compact we may take R sufficiently large such that $K_\varepsilon \subseteq \mathbb{R}^n \times [-R,R]$; then one has

$$\mathbb{P}_x\left(|\zeta_t^\zeta| \leq R\right) \geq p_t^x(K_\varepsilon) \geq 1-\varepsilon, \quad \text{for all } t \geq 0. \tag{5.13}$$

However $\zeta_t^\zeta \to \infty$ so we may take t sufficiently large that $|\zeta_t^\zeta| > R$. This contradicts (5.13), hence p_t^x is not tight. □

Example 5.1.9 (UFG-Grušin Plane)**.** We give here a simple example of a process that satisfies the Obtuse Angle Condition but is not tight. Let $d = 1$, $N = 2$ and

$$V_0 = k\zeta\partial_\zeta, \quad V_1 = \zeta\partial_z, \quad k \in \mathbb{R}\,.$$

This corresponds to the SDE

$$d\zeta_t = k\zeta_t dt$$

$$dZ_t = \sqrt{2}\zeta_t \circ dB_t,$$

where $\{B_t\}_{t\geq 0}$ is a one-dimensional Brownian motion. Because $[V_1,V_0] = -kV_1$, we have

$$([V_1,V_0]f)(V_1 f) = -k(V_1 f)^2$$

therefore the Obtuse Angle Condition, (2.8) is satisfied if and only if $k > 0$ (it is also shown in [32, Example 4.4] that, if $k > 0$, then $V_1(\mathcal{P}_t f)(\cdot)$ decays exponentially fast with rate $-2k$). On the other hand, if $k > 0$ the process is not tight. Indeed, Hypothesis 5.1.1 ([**H.2**]) is satisfied if and only if $k < 0$, as we come to show. To this end, we first we solve the SDE, and find

$$\zeta_t = \zeta e^{kt}$$

$$Z_t = Z_0 + \sqrt{2}\zeta \int_0^t e^{ks} \circ dB_s.$$

As a consequence of Proposition 5.1.8, the whole process (Z_t, ζ_t) is not tight if $k > 0$. However in this case also the process Z_t, seen as a non-autonomous one dimensional SDE, is not tight when $k > 0$. Indeed suppose by contradiction that ([**H.2**]) holds; then for any $\varepsilon > 0$ there exists $R > 0$ such that

$$Q_{0,t}\mathbb{1}_{[-R,R]}(z) \geq 1 - \varepsilon\,, \quad \text{for all } t \geq 0. \tag{5.14}$$

However if $Z_0 = z$ then Z_t is normally distributed with mean z and variance $\zeta^2(e^{2kt} - 1)/k$, so we may write

$$Z_t = z + \zeta\sqrt{\frac{e^{2kt} - 1}{k}}\xi \tag{5.15}$$

where ξ is a one-dimensional standard normal random variable. Then we have

$$Q^\zeta_{0,t}\mathbb{1}_{[-R,R]}(z) = \mathbb{E}\mathbb{1}_{[-R,R]}(Z_t^{0,z,\zeta}) = \mathbb{P}\left(\left|z + \zeta\sqrt{\frac{e^{2kt} - 1}{k}}\xi\right| \leq R\right) \to 0 \quad \text{as } t \to \infty,$$

which contradicts (5.14). Note that if $k = 0$ then $Z_t = \sqrt{2}\zeta B_t$ which is not tight by a similar argument. However if $k < 0$ then the process Z_t is tight. Indeed, assume that $k = -\ell < 0$; to see that $\{q_t^{0,z}\}_{t\geq 0}$ is a tight family of measures, it is sufficient to apply a Lyapunov criterion and show that the function $\varphi(z) = z^2$ satisfies $\sup_t Q_{0,t}\varphi(y) < \infty$ (when $k = -\ell < 0$). To prove the latter fact, observe that if $(Z_s, \zeta_0) = (z, \zeta)$ then, by (5.15), we get

$$Q^\zeta_{s,t}\varphi(z) = \mathbb{E}[Z_t^2|\zeta_0 = \zeta, Z_s = z] = z^2 + \frac{\zeta^2 e^{-2\ell s}}{\ell}(1 - e^{-2\ell(t-s)}) \leq z^2 + \frac{e^{-2\ell s}}{\ell}\zeta^2.$$

If $k < 0$ we see that $X_t = (Z_t, \zeta_t)$ converges in distribution. □

Example 5.1.10. We conclude this section with an example which satisfies all the points in Hypothesis 5.1.1 in a non-trivial way, in the sense that it exhibits many invariant measures. Take $k > 1$ and consider the following SDE

$$d\zeta_t = -\sin(\zeta_t)dt$$

$$dZ_t = -kZ_t dt + \sqrt{2}\zeta_t \circ dB_t.$$

In this case

$$V_0 = -\sin(\zeta)\partial_\zeta - kz\partial_z, \quad V_1 = \zeta\partial_z, U_0 = -kz\partial_z, U_1 = \zeta\partial_z.$$

Then we have

$$[V_1, V_0] = [\zeta\partial_z, -\sin(\zeta)\partial_\zeta - kz\partial_z] = -k\zeta\partial_z - \sin(\zeta)\partial_z = \left(-k + \frac{\sin(\zeta)}{\zeta}\right) V_1.$$

Note that the function $h(\zeta) = \sin(\zeta)/\zeta$ is bounded and smooth, when extended to the origin with the value $h(0) = 1$, so the UFG condition is satisfied at level $m = 1$. Moreover,

$$([V_1, V_0]f)(V_1 f) = -(k + \frac{\sin(\zeta)}{\zeta}) |V_1 f|^2 \leq -(k-1) |V_1 f|^2$$

and hence (2.8) is satisfied. To apply the results of Section 5.1 we must show that Hypothesis 5.1.1 holds. Note that the vector field V_1 is non-zero except when $\zeta = 0$ therefore Hypothesis 5.1.1 **[H.1]** is satisfied everywhere that $\zeta \neq 0$. To show that Hypothesis 5.1.1 **[H.2]** holds we consider a function $\varphi \in C^2(\mathbb{R})$ such that $\varphi(z) = |z|$ for $|z| > 1$. Then, for $|z| > 1$, one has

$$\mathcal{L}_t\varphi(z) = -kz\varphi'(z) + \zeta_t^2\varphi''(z) = -kz\mathrm{sign}(z) = -k\varphi(z).$$

Therefore φ is a Lyapunov function so by Note 5.1.2 we have that the measures $\{q_t^{0,z} : t \geq 0\}$ are tight for any $z \in \mathbb{R}$ and Hypothesis 5.1.1 **[H.2]** is satisfied. We also have that ζ_t converges for any $\zeta \in \mathbb{R}$ and the limit $\overline{\zeta}$ is given by

$$\overline{\zeta} = \begin{cases} 2n\pi & \text{for } \zeta \in ((2n-1)\pi, (2n+1)\pi) \text{ for some } n \in \mathbb{Z} \setminus \{0\} \\ (2n+1)\pi & \text{for } \zeta = (2n+1)\pi \text{ for some } n \in \mathbb{Z} \\ 0 & \text{for } \zeta \in (-\pi, \pi). \end{cases}$$

Hence for $\zeta \notin (-\pi, \pi)$ we may apply Theorem 5.1.6 to obtain that $X_t = (Z_t, \zeta_t)$ converges in distribution to a random variable which is distributed according to the unique invariant measure defined on the line $\mathbb{R} \times \{\overline{\zeta}(\zeta_0)\}$. Moreover, for $\zeta = n\pi$ for some $n \in \mathbb{Z} \setminus \{0\}$ we see that $\zeta_t = \zeta$ and Z_t satisfies the Ornstein Uhlenbeck SDE

$$dZ_t = -kZ_t dt + \sqrt{2}\zeta dB_t.$$

In particular, in this case Z_t has a unique invariant measure and this is given by a Gaussian measure with mean 0 and variance ζ^2/k. Therefore for any $n \in \mathbb{Z}\setminus\{0\}$ and $\zeta \in ((2n-1)\pi, (2n+1)\pi)$ we have that X_t converges in distribution to $(\frac{2n\pi}{\sqrt{k}}\xi, 2n\pi)$, where ξ is a one-dimensional standard normal random variable.[21] $\square$

5.2 Long-time behaviour of UFG diffusions: general case

In the previous section we investigated the case in which the representation of the form "ODE+SDE" is global. In this section we study the general UFG-case, in which such a representation is, in general, only local. That is, we finally address the full problem of analysing the asymptotic behaviour of (1.1), assuming that the vector fields $V_0, \ldots, V_d$ satisfy the UFG condition (see Definition 2.2.1). This case is substantially richer than the one considered in Section 5.1; however the fact that, locally, we can always represent the SDE (1.1) as a system of the form "ODE+SDE", still means that we should be able to identify a suitable ODE which drives the dynamics. We will demonstrate that this is indeed the case and that such an ODE is the integral curve of the vector field $V_0^{(\perp)}$; that is, the curve

$$\zeta_t = e^{tV_0^{(\perp)}} x_0 \,, \tag{5.16}$$

where x_0 is the initial datum of the SDE (1.1), i.e. $X_0 = x_0$. This should not be a surprise in view of Proposition 3.2.3. Nevertheless, to understand why this is the case, it is useful to build an analogy with the setting of the previous section:

[21] Since Z_t satisfies a non-autonomous Ornstein Uhlenbeck equation one can also study its asymptotic behaviour more directly, see e.g. [63].

if the SDE is of the form (5.1)-(5.2), then $V_0^{(\perp)} = (0, \ldots, 0, W_0)$. Therefore in the simplified setting (5.1)-(5.2), the ODE (5.16) substantially reduces to (5.2). The previous sentence is correct for less than observing that (5.16) is an N-dimensional ODE, while (5.2) is a one-dimensional curve. We keep using the notation ζ_t for both curves only to emphasize the analogy; however, while the one-dimensional autonomous nature of the ODE (5.2) implies that its solution has a limit, the zoology of possible behaviours for the curve (5.16) is much more varied. In this thesis we only analyse the case in which the curve (5.16) converges to a limit and in future work we will treat the case when (5.2) does not have a solution. However, roughly speaking, in Theorem 5.3.10, we prove that a necessary condition for the SDE (1.1) to have an invariant measure is that the ODE (5.16) should admit one as well (notice that if the curve (5.16) converges to a limit point $\bar{x}$, then it admits the Dirac measure $\delta_{\bar{x}}$ as invariant measure).

As anticipated in the introduction, the above discussion motivates introducing the process

$$\mathscr{Z}_t := e^{-tV_0^{(\perp)}}(X_t^{(x_0)}). \tag{5.17}$$

Clearly $\mathscr{Z}_0 = x_0$, so $\mathscr{Z}_t$ and $X_t^{(x_0)}$ start from the same point. This process is time-inhomogeneous (as we show at the beginning of Section 5.3.1) and it will have a central role in what follows, hence further comments on the definition (5.17) are in order:

- To continue drawing the useful parallel with Section 5.1, notice that this process plays in this context an analogous role to the one that Z_t (solution of (5.1)) has in Section 5.1, see Note 5.1.3.

- Let us recall that if $X_0 \in S_{x_0}$ then $X_t \in \overline{\mathscr{S}}_{x_0}$ for every $t \geq 0$ (see Proposition 3.1.7); more precisely, for every $t \geq 0$ X_t belongs to the integral submanifold $\overline{S}_{e^{tV_0^{(\perp)}}(x_0)}$ almost surely (see Proposition 3.2.3). We will make assumptions to guarantee that X_t hits neither the boundary of $S_{e^{tV_0^{(\perp)}}x_0}$ nor the boundary of $\mathscr{S}_{x_0}$ *in finite time* (see Hypothesis 5.3.3 [**A.4**], Lemma 3.1.16 and Note 5.3.4 for more precise comments on this). Therefore $\mathscr{Z}_t$ lives on the manifold S_{x_0}, for every $t \geq 0$. So, in the end, while X_t takes values in $\mathscr{S}_{x_0}$, $\mathscr{Z}_t$ takes values in $S_{x_0} \subseteq \mathscr{S}_{x_0}$. One can informally think of $\mathscr{Z}_t$ as being a "projection" of X_t

on the submanifold $S_{x_0} \subseteq \mathscr{S}_{x_0}$, see again Note 5.1.3.

- Finally, on a small technical point, as we have already observed in Note 3.1.15, $V_0^{(\perp)}$ may not be uniformly Lipschitz. However, to avoid problems of well posedness and uniqueness, throughout this section we assume that $V_0^{(\perp)}$ is indeed Lipschitz.

We will show that the time-inhomogeneous process $\mathcal{Z}_t$ can be studied by means of slight modifications of the approach used in Section 5.1 to study the process (5.1). Therefore the strategy (and one of the the main novelties) of this section is to use the auxiliary time-inhomogeneous process $\mathcal{Z}_t$ in order to make deductions on the behaviour of the time-homogeneous process X_t. We carry out this programme in Section 5.3.1 below. Before moving on, we give a simple example which demonstrates that $\mathcal{Z}_t \in S_{x_0}$ for every $t \geq 0$ and, in Section 5.3, we gather further preliminary results on the process $\mathcal{Z}_t$.

Example 5.2.1 (Random Circles continued)**.** Consider again Example 3.1.11, in the case in which the initial datum is $(x_0, y_0) = (1, 0)$. Using (3.16) and (3.17)-(3.18), we have

$$
\begin{aligned}
\mathcal{Z}_t := e^{-tV_0^{(\perp)}}(X_t, Y_t) &= \begin{pmatrix} X_t \cos(-t) - Y_t \sin(-t) \\ X_t \sin(-t) + Y_t \cos(-t) \end{pmatrix} \\
&= \begin{pmatrix} e^{\sqrt{2}B_t} \cos(t)\cos(t) + e^{\sqrt{2}B_t} \sin(t)\sin(t) \\ -e^{\sqrt{2}B_t} \cos(t)\sin(t) + e^{\sqrt{2}B_t} \sin(t)\cos(t) \end{pmatrix} \\
&= e^{\sqrt{2}B_t} \begin{pmatrix} 1 \\ 0 \end{pmatrix}
\end{aligned}
$$

In particular, $\mathcal{Z}_t$ takes values in the positive half-line, which is precisely $S_{(1,0)} = S_{(x_0,y_0)}$. $\square$

5.3 The auxiliary process $\mathcal{Z}_t$ and its associated two-parameter semigroup

By differentiating (5.17) we see that $\mathcal{Z}_t$ satisfies the following SDE

$$
\begin{aligned}
d\mathcal{Z}_t &= -V_0^{(\perp)}(e^{-tV_0^{(\perp)}}(X_t))dt + \left(\mathcal{J}_x e^{-tV_0^{(\perp)}}\right)(X_t)V_0(X_t)dt \\
&\quad + \sqrt{2}\sum_{i=1}^{d}\left(\mathcal{J}_x e^{-tV_0^{(\perp)}}\right)(X_t)V_i(X_t)\circ dB_t^i \\
&= -V_0^{(\perp)}(\mathcal{Z}_t)dt + \mathrm{Ad}_{tV_0^{(\perp)}}V_0(\mathcal{Z}_t)dt + \sqrt{2}\sum_{i=1}^{d}\mathrm{Ad}_{tV_0^{(\perp)}}V_i(\mathcal{Z}_t)\circ dB_t^i,
\end{aligned}
$$

where, as customary, we have set $(\mathrm{Ad}_{tV}Y)(x) := (\mathcal{J}_x e^{-tV})(e^{tV}(x))\cdot Y(e^{tV}x)$, for any two smooth vector fields V and Y. By using (1.12), the elementary property $\mathrm{Ad}_{tV}V = V$ and introducing the notation

$$
\begin{aligned}
\mathcal{V}_{0,t} &:= \mathrm{Ad}_{tV_0^{(\perp)}}V_0^{(\hat{\Delta})} \\
\mathcal{V}_{j,t} &:= \mathrm{Ad}_{tV_0^{(\perp)}}V_j \quad j\in\{1,\dots,d\},
\end{aligned}
\tag{5.18}
$$

we conclude that $\mathcal{Z}_t$ satisfies the following SDE with time-dependent coefficients:

$$
\begin{aligned}
d\mathcal{Z}_t &= \mathrm{Ad}_{tV_0^{(\perp)}}V_0^{(\hat{\Delta})}(\mathcal{Z}_t)dt + \sqrt{2}\sum_{i=1}^{d}\mathrm{Ad}_{tV_0^{(\perp)}}V_i(\mathcal{Z}_t)\circ dB_t^i \\
&= \mathcal{V}_{0,t}(\mathcal{Z}_t)dt + \sqrt{2}\sum_{i=1}^{d}\mathcal{V}_{i,t}(\mathcal{Z}_t)\circ dB_t^i\,,
\end{aligned}
\tag{5.19}
$$

As usual, we denote by $\mathcal{P}_t$ the one parameter semigroup associated with X_t; the two-parameter semigroup associated with $\mathcal{Z}_t$ is instead given by

$$
\mathcal{Q}_{s,t}f(z) = \mathbb{E}\left[f(\mathcal{Z}_t)|\mathcal{Z}_s = z\right], \quad z\in S_{x_0}, s\le t, f\in C_b(\mathbb{R}^N).
$$

The semigroups $\mathcal{Q}_{s,t}$ and $\mathcal{P}_t$ are related as follows:

$$\mathcal{P}_t f(x) = \mathcal{Q}_{s,s+t}(f \circ e^{(s+t)V_0^{(\perp)}})(e^{-sV_0^{(\perp)}}(x)), \quad x \in S_{e^{sV_0^{(\perp)}}(x_0)}, \; s \in \mathbb{R}, t \geq 0, \; f \in C_b(\mathbb{R}^N),$$

$$\mathcal{Q}_{s,t} g(z) = \mathcal{P}_{t-s}(g \circ e^{-tV_0^{(\perp)}})(e^{sV_0^{(\perp)}}(z)), \qquad z \in S_{x_0}, \; 0 \leq s \leq t, \; g \in C_b(\overline{S}_{x_0}) \,. \tag{5.20}$$

We stress that $\{\mathcal{Q}_{s,t}\}_{0 \leq s \leq t}$ is defined on S_{x_0} (as per Hypothesis 5.3.3 below). In (5.20) we consider functions g which are continuous up to and including the boundary of S_{x_0} for purely technical reasons (see proof of Proposition 5.3.2).
In Proposition 5.3.2 we make some clarifications on the smoothing properties of the semigroup $\mathcal{Q}_{s,t}$. To state such a lemma, we need to properly formulate some preliminary facts. Consider the following "hierarchy" of operators:

$$\mathcal{V}_{[i],t} := \mathcal{V}_{i,t} \qquad i = 0, 1, \ldots, d \text{ (defined as in (5.18))}$$

$$\mathcal{V}_{[\alpha * 0],t} := [\mathcal{V}_{[\alpha],t}, \mathcal{V}_{[0],t} + \partial_t], \qquad \alpha \in \mathcal{A},$$

$$\mathcal{V}_{[\alpha * i],t} := [\mathcal{V}_{[\alpha],t}, \mathcal{V}_{[i],t}], \qquad \alpha \in \mathcal{A}, i = 1, \ldots, d \,.$$

For each $\alpha \in \mathcal{A}$ we can view the vector field $(z, t) \mapsto \mathcal{V}_{[\alpha],t}(z)$ as a vector field on $\mathbb{R}^N$, the coefficients of which depend on time or as a vector field on $\mathbb{R}^N \times \mathbb{R}$. We can define the UFG condition for vector fields in $\mathbb{R}^N \times \mathbb{R}$ in an analogous way to Definition 2.2.1. In Proposition 5.3.1 we prove that the set of vector fields $\{\mathcal{V}_{[0],t} + \partial_t, \mathcal{V}_{[1],t}, \ldots, \mathcal{V}_{[d],t}\}$ satisfy the UFG condition on $\mathbb{R}^N \times \mathbb{R}$ provided the vector fields $\{V_0, V_1, \ldots, V_d\}$ satisfy the UFG condition on $\mathbb{R}^N$.

Proposition 5.3.1. *Assume that the vector fields $\{V_0, V_1, \ldots, V_d\}$ on $\mathbb{R}^N$ satisfy the UFG condition at level m; then the vector fields $\{\partial_t + \mathcal{V}_{[0],t}, \mathcal{V}_{[1],t}, \ldots, \mathcal{V}_{[d],t}\}$ satisfy the UFG condition at level m when viewed as vector fields on $\mathbb{R}^N \times \mathbb{R}$. Moreover, for any $\alpha \in \mathcal{A}_m$,*

$$\mathcal{V}_{[\alpha],t} = \mathrm{Ad}_{tV_0^{(\perp)}} V_{[\alpha]}. \tag{5.21}$$

Proof of Proposition 5.3.1. The proof is deferred to Appendix A.2.4. $\square$

Recall from Section 2.2 that the map $z \in S_{x_0} \mapsto \mathcal{P}_t f(z)$ is smooth (along the directions $V_{[\alpha]}$, $\alpha \in \mathcal{A}_m$) for any $f \in C_b(\mathbb{R}^N)$. In Proposition 5.3.2 we show that for each fixed $s < t$ the map $z \in S_{x_0} \mapsto \mathcal{Q}_{s,t}g(z)$ is also smooth in the directions $\mathcal{V}_{[\alpha],s}$ for any $g \in C_b(\overline{S}_{x_0})$ and $\alpha \in \mathcal{A}_m$. A key observation to understand the statement of Proposition 5.3.2 is the following one:

$$V \in \hat{\Delta} \text{ and } \hat{\Delta} \text{ is invariant under the vector field } W \quad \Rightarrow \quad \mathrm{Ad}_{tW}V \in \hat{\Delta}.^{22} \tag{5.22}$$

In particular, $\mathcal{V}_{j,t} \in \hat{\Delta}$ for every $j \in \{0, \ldots, d\}$.

Proposition 5.3.2. *Assume the vector fields $\{V_0, \ldots, V_d\}$ satisfy the UFG condition and that the vector $V_0^{(\perp)}$ is uniformly Lipschitz. Then, for any $g \in C_b(\overline{S}_{x_0})$, the map $(z, s) \mapsto \mathcal{Q}_{s,t}g(z)$ is differentiable in the time variable s and in the spatial directions $\mathcal{V}_{[\alpha],s}$ for any $z \in \overline{S}_{x_0}, t > s, \alpha \in \mathcal{A}_m$. Moreover $\mathcal{Q}_{s,t}g(z)$ satisfies the equation*

$$\partial_s \mathcal{Q}_{s,t}g(z) = -\mathcal{L}_s \mathcal{Q}_{s,t}g(z), \qquad \textit{for any } z \in S_{x_0}, s < t. \tag{5.23}$$

Here $\mathcal{L}_s$ is the differential operator defined as

$$\mathcal{L}_s\psi(z) = \mathcal{V}_{0,s}\psi(z) + \sum_{i=1}^{d} \mathcal{V}_{i,s}^2 \psi(z)\,,$$

for $\psi : \overline{S}_{x_0} \to \mathbb{R}$ sufficiently smooth.

Proof of Proposition 5.3.2. The proof is deferred to Appendix A.2.4. □

5.3.1 Convergence to Equilibria

We now turn to the asymptotic behaviour of the process $\mathcal{Z}_t$. As we have already stated, we will concentrate on the case in which the solution of the ODE (5.16) converges. Let us define the map

$$\begin{aligned} W^\infty : \mathrm{Dom}(W^\infty) \subseteq \mathbb{R}^N &\longrightarrow \mathbb{R}^N \\ x &\longrightarrow \lim_{t\to\infty} e^{tV_0^{(\perp)}}(x). \end{aligned}$$

[22]Indeed, by the definition of invariance (see Definition 2.3.1), we have that $\mathcal{J}_x e^{tW}(x)$ maps $\hat{\Delta}(x)$ to $\hat{\Delta}(e^{tW}(x))$. Therefore $\mathcal{J}_x e^{tW}(e^{-tW}(x))$ maps $\hat{\Delta}(e^{-tW}(x))$ to $\hat{\Delta}(x)$. Now $V \in \hat{\Delta}$, so $V(e^{-tW}(x)) \in \hat{\Delta}(e^{-tW}(x))$ and we have that $\mathrm{Ad}_{tW}V(x) = \mathcal{J}_x e^{tW}(x)V(e^{-tW}(x)) \in \hat{\Delta}(x)$. That is, $\mathrm{Ad}_{tW}V \in \hat{\Delta}$.

Here $\mathrm{Dom}(W^\infty)$ is the set of all points $x \in \mathbb{R}^N$ such that the integral curve $e^{tV_0^{(\perp)}}(x)$ converges to a finite limit as t tends to ∞.

Hypothesis 5.3.3. Assume the following:

[A.1] The vector fields $\{V_0, V_1, \dots, V_d\}$ satisfy the UFG condition.

[A.2] The vector field $V_0^{(\perp)}$ is uniformly Lipschitz.

[A.3] Define the measures p_t^x by $p_t^x(A) = \mathcal{P}_t \mathbb{1}_A(x)$ for any Borel measurable $A \subseteq \mathbb{R}^N$. The family $\{p_t^x : t \geq 0\}$ is tight for all $x \in \mathbb{R}^N$.

[A.4] Define the measures $\mathfrak{q}_t^{s,z}$ by $\mathfrak{q}_t^{s,z}(A) = \mathcal{Q}_{s,t}\mathbb{1}_A(z)$ for any Borel measurable $A \subseteq S_{x_0}$. Then we require that for each $z \in S_{x_0}$ the measures $\{\mathfrak{q}_t^{0,z} : t \geq 0\}$ are tight on S_{x_0}; that is, for all $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subseteq S_{x_0}$ such that $\mathfrak{q}_t^{0,z}(K_\varepsilon) \geq 1 - \varepsilon$ for all $t \geq 0$.

[A.5] The long time derivative estimates (2.9) and (2.11) are satisfied.

[A.6] The initial datum x_0 of the SDE (1.1) is such that the curve (5.16) started at x_0, admits a limit, i.e. there exists $\bar{x} \in \mathbb{R}^N$ such that $V_0^{(\perp)}(\bar{x}) = 0$ and $e^{tV_0^{(\perp)}}(x_0) \to \bar{x}$ as $t \to \infty$.

[A.7] Assumptions on the map W^∞: the domain of W^∞ contains the whole manifold $\overline{S}_{x_0}$ and the image of S_{x_0} through W^∞ is all contained in a submanifold of $\hat{\Delta}$. More explicitly, there exists an integral submanifold of $\hat{\Delta}$, $S_{\overline{x}}$, such that $V_0^{(\perp)} = 0$ on $S_{\overline{x}}$ and the image of S_{x_0} through W^∞ is all contained in $S_{\overline{x}}$, $W^\infty(S_{x_0}) \subseteq S_{\overline{x}}$. Furthermore we assume that W^∞ is a continuous map from $\overline{S}_{x_0} \cup \mathscr{S}_{x_0} \cup S_{\overline{x}}$ into $\mathbb{R}^N$.

Note 5.3.4. Some comments on the above assumptions, in the order in which they are stated.

- As a general premise, observe that, for every fixed $t \geq 0$, $X_t \in \mathscr{S}_{x_0}$ if and only if $\mathcal{Z}_t \in S_{x_0}$. Indeed, $X_t = e^{tV_0^{(\perp)}}\mathcal{Z}_t$ so if $\mathcal{Z}_t$ is in S_{x_0} then in particular it is in $\mathscr{S}_{x_0}$ and X_t is just obtained by moving along an integral curve of $V_0^{(\perp)}$; hence, by construction of the manifold $\mathscr{S}_{x_0}$, X_t is still in $\mathscr{S}_{x_0}$. The validity of the reverse implication can be argued similarly (using Lemma 3.1.16 and

Proposition 3.2.3 as well). As a consequence, if $\mathcal{Z}_t$ doesn't hit the boundary of S_{x_0} *in finite time* then X_t doesn't hit the boundary of $\mathscr{S}_{x_0}$ *in finite time.*

- Hypothesis 5.3.3 [**A.4**] implies that $\mathcal{Z}_t \in S_{x_0}$ almost surely, for every $t \geq 0$, i.e. it implies that $\mathcal{Z}_t$ doesn't hit the boundary of S_{x_0} in finite time. Indeed, assume by contradiction that there exists $t_0 > 0$ such that $\mathbb{P}(\mathcal{Z}_{t_0} \in \partial S_{x_0}) =: \varepsilon > 0$. Recall $\partial S_{x_0} := \overline{S}_{x_0} \setminus S_{x_0}$. By the previous bullet point if $\mathcal{Z}_{t_0}$ belongs to ∂S_{x_0} then $X_{t_0} \in \partial \mathscr{S}_{x_0}$. By Proposition 3.2.1 we then have that X_t is in the boundary of $\mathscr{S}_{x_0}$ for any $t > t_0$. That is,

$$\mathbb{P}(X_t \in \partial \mathscr{S}_{x_0}) \geq \mathbb{P}(X_{t_0} \in \partial \mathscr{S}_{x_0}) \geq \mathbb{P}(\mathcal{Z}_{t_0} \in \partial S_{x_0}) = \varepsilon > 0\,, \quad \text{for any } t > t_0.$$ [23]

We know from [**A.3**] that, given ε as in the above, there exists a compact set $K_{\varepsilon/2} \subseteq S_{x_0}$ such that $\mathbb{P}(\mathcal{Z}_t \in K_{\epsilon/2}) = \mathfrak{q}_t(K_{\varepsilon/2}) \geq 1 - \varepsilon/2$ for every $t \geq 0$. Now using that $\mathscr{S}_{x_0}$ and $\partial \mathscr{S}_{x_0} = \overline{\mathscr{S}}_{x_0} \setminus \mathscr{S}_{x_0}$ are disjoint, for every $t > t_0$ we have

$$\begin{aligned} 1 &= \mathbb{P}(X_t \in \partial \mathscr{S}_{x_0}) + \mathbb{P}(X_t \in \mathscr{S}_{x_0}) \\ &\geq \mathbb{P}(\mathcal{Z}_{t_0} \in \partial S_{x_0}) + \mathbb{P}(\mathcal{Z}_t \in K_{\varepsilon/2}) + \mathbb{P}(\mathcal{Z}_t \in (K_{\varepsilon/2})^C) \\ &\geq \mathbb{P}(\mathcal{Z}_{t_0} \in \partial S_{x_0}) + \mathbb{P}(\mathcal{Z}_t \in K_{\varepsilon/2}) \geq \varepsilon + 1 - \varepsilon/2 = 1 + \varepsilon/2\,, \end{aligned}$$

where in the first inequality we have used the observation in the first bullet point of this note and $(K_{\varepsilon/2})^C$ denotes complement in $\mathscr{S}_{x_0}$. Hence $\varepsilon = 0$, i.e. $\mathcal{Z}_t$ belongs to S_{x_0} almost surely.

- Hypothesis 5.3.3 [**A.6**] is the analogous of Hypothesis 5.1.1. [**H.4**].

- Hypothesis 5.3.3 [**A.7**] is slightly more complicated to explain, so we observe that it is satisfied in the representation of the form "ODE+SDE" (5.1)-(5.3) of the previous section, if $\zeta_t = e^{tW_0}\zeta_0$ converges to some $\bar{\zeta}$. Indeed in that case if $x_0 = (z_0, \zeta_0)$ then $S_{x_0} = \mathcal{H}_{\zeta_0}$ and $S_{\overline{x}} = \mathcal{H}_{\bar{\zeta}}$ (both of these manifolds are n-dimensional hyperplanes in $\mathbb{R}^{n+1}$, hence they are closed). Moreover, for every $x = (z, \zeta) \in \mathscr{S}_{x_0}$, $W^\infty(x) = W^\infty((z, \zeta)) = (z, \bar{\zeta})$, hence the map W^∞ is

[23] The second inequality is an inequality rather than an equality because of Lemma 3.1.16.

continuous on $\mathscr{S}_{x_0}$. If $x = (z, \bar{\zeta}) \in S_{\overline{x}}$ then $W^\infty(x) = x$, so W^∞ is continuous on $S_{\overline{x}}$ as well. Because in this case the map W^∞ is just a projection on the plane $S_{\overline{x}}$, W^∞ is continuous on $S_{x_0} \cup \mathscr{S}_{x_0} \cup S_{\overline{x}} = \mathscr{S}_{x_0} \cup S_{\overline{x}}$ (the equality holding because $S_{x_0} \subset \mathscr{S}_{x_0}$).

- By [**A.6**] $V_0^{(\perp)}(\bar{x}) = 0$; using Lemma 3.1.1, this implies that $S_{\bar{x}} = \mathscr{S}_{\bar{x}}$. Hence, by Lemma 3.1.17, $V_0^{(\perp)}(x) = 0$ for every $x \in S_{\bar{x}}$. So in reality [**A.6**] implies that part of [**A.7**] where we require $V_0^{(\perp)}$ to vanish on the whole $S_{\bar{x}}$.

- If we don't make any assumptions on the map W^∞, when we look at the set $W^\infty(S_{x_0})$, it may occur that this is not a connected set and, even if it were connected, it may be contained in more than one submanifold of $\hat{\Delta}$ (see Example 5.1.9). If we assume that W^∞ is continuous, because S_{x_0} is connected then also $W^\infty(S_{x_0})$ is; for simplicity, we are also explicitly assuming that $W^\infty(S_{x_0})$ is contained in just one submanifold of $\hat{\Delta}$, the manifold $S_{\overline{x}}$. It could also occur that on the limit manifold $W^\infty(S_{x_0})$ we have that $V_0^{(\perp)}(x) \neq 0$ for every $x \in W^\infty(S_{x_0})$, see for instance Example 5.1.10. If this is the case, then one can take such a manifold as starting manifold and apply the theory that we explain here by taking starting points on this manifold; i.e. one can sort of "repeat the procedure" illustrated here by starting the dynamics again on that manifold. So, in conclusion one just needs to study the case in which $V_0^{(\perp)}(x) = 0$ for every $x \in W^\infty(S_{x_0})$. Again for simplicity, we assume $V_0^{(\perp)}(x) \neq 0$ for every $x \in S_{\overline{x}}$.

- Finally, notice that if W^∞ is well defined and continuous on S_{x_0} then W^∞ is also a well-defined and continuous map from $\mathscr{S}_{x_0}$ to $\mathbb{R}^N$. We show this fact in Lemma A.1.17 , contained in Appendix A.1.5. Notice also that W^∞ is the identity when restricted to $S_{\overline{x}}$, hence W^∞ is always well defined and continuous on $S_{\overline{x}}$. What we are requiring with the last point of Hypothesis 5.3.3 is that the map should be continuous not only on each one of the manifolds $\overline{S}_{x_0}, \mathscr{S}_{x_0}$ and $S_{\overline{x}}$, but also that it should be continuous on the union of these three sets. The reason why we need continuity also on the closure of S_{x_0} is, again, technical, see proof of Lemma A.2.11

Before we consider the behaviour of $X_t^{(x)}$ in the case when $e^{tV_0^{(\perp)}}(x)$ is convergent, we must first consider the trivial case, i.e. the behaviour of the process when we start it from the "equilibrium manifold" $S_{\overline{x}}$, where $V_0^{(\perp)}(x) = 0$. We do this in Proposition 5.3.5 below, which is the analogous of Lemma 5.1.4.

Proposition 5.3.5. *Let Hypothesis 5.3.3* [**A.1**], [**A.3**] *and* [**A.5**] *hold. Let S be an integral submanifold of $\hat{\Delta}$ such that $V_0^{(\perp)} = 0$ on S. Then there exists a unique invariant measure $\overline{\mu}^S$ of $\mathcal{P}_t$ supported on $\overline{S}$ such that*

$$\lim_{t\to\infty} \mathcal{P}_t f(x) = \overline{\mu}^S(f), \text{ for all } x \in S, f \in C_b(\mathbb{R}^N). \tag{5.24}$$

Moreover the convergence is uniform on compact subsets of S; that is, for every compact set $K \subseteq S$ and every $f \in C_b(\mathbb{R}^N)$ we have

$$\lim_{t\to\infty} \sup_{x\in K} \left|\mathcal{P}_t f(x) - \overline{\mu}^S(f)\right| = 0\,.$$

Proof of Proposition 5.3.5. The proof is deferred to Appendix A.2.4 . □

Note 5.3.6. The assumption that $V_0^{(\perp)} = 0$ on S implies that, if $x \in S$, then the map $t \mapsto \mathcal{P}_t f(x)$ is differentiable, for any $f \in C_b(\mathbb{R}^N)$. Indeed, as explained in the Introduction, in general we have that $\mathcal{P}_t f$ is differentiable in the direction $\partial_t - V_0$ and in the directions contained in $\hat{\Delta}$ (see Appendix A.1.4) and satisfies

$$(\partial_t - V_0)\mathcal{P}_t f = \sum_{i=1}^{d} V_i^2 \mathcal{P}_t f.$$

However if $V_0^{(\perp)}(x) = 0$ for all $x \in S$ then $V_0(x) \in \hat{\Delta}(x)$ for all $x \in S$ and hence $\mathcal{P}_t f$ is also differentiable in the direction V_0 on S. Therefore we have that $\mathcal{P}_t$ is also differentiable in time, i.e. as a map $t \mapsto \mathcal{P}_t f$, and satisfies

$$\partial_t \mathcal{P}_t f = V_0 \mathcal{P}_t f + \sum_{i=1}^{d} V_i^2 \mathcal{P}_t f.$$

□

By Hypothesis [**A.7**] $V_0^{(\perp)} = 0$ on $S_{\overline{x}}$ so we can apply Proposition 5.3.5 to the manifold $S_{\overline{x}}$ and throughout the rest of the section we shall denote by $\overline{\mu}^{S_{\overline{x}}}$ the invariant measure supported on $\overline{S}_{\overline{x}}$ such that (5.24) holds for all $x \in S_{\overline{x}}$. Such a measure exists and is unique by Proposition 5.3.5. Similarly to what we did in Section 5.1, equation (5.11), we shall extend this to a measure $\mu^{S_{\overline{x}}}$ defined on $\mathbb{R}^N$ by setting

$$\mu^{S_{\overline{x}}}(A) = \overline{\mu}^{S_{\overline{x}}}(A \cap \overline{S}_{\overline{x}}), \text{ for any Borel measurable set } A \subseteq \mathbb{R}^N.$$

For any $\overline{x} \in \mathbb{R}^N$, let $\mathcal{I}_0(\overline{x}) = \{x \in \mathbb{R}^N : W^\infty(S_x) \subseteq S_{\overline{x}}\}$. The set $\mathcal{I}_0$ is contained within the basin of attraction for the measure $\mu^{S_{\overline{x}}}$. Indeed, Theorem 5.1.5 below shows that for all $x \in \mathcal{I}_0(\overline{x})$ we have that $\mathcal{P}_t f(x)$ converges to $\mu^{S_{\overline{x}}}(f)$, for all $f \in C_b(\mathbb{R}^N)$.

Theorem 5.3.7. *Let Hypothesis 5.3.3 hold. Let $\overline{x} \in \mathbb{R}^N$ be such that $V_0^{(\perp)}(\overline{x}) = 0$. Then there exists an invariant measure $\mu^{S_{\overline{x}}}$ supported on $\overline{S}_{\overline{x}}$ such that for each $x_0 \in \mathcal{I}_0(\overline{x})$, and $f \in C_b(\mathbb{R}^N)$ we have that $\mathcal{P}_t f(x_0)$ converges to $\mu^{S_{\overline{x}}}(f)$.*

Proof. Throughout the proof we fix an arbitrary point $x_0 \in \mathcal{I}_0$. The proof is split into 3 steps.

- *Step 1*: We first construct a tight evolution system of measures, $\{\nu_t\}_{t\geq 0}$, for the semigroup $\{\mathcal{Q}_{s,t}\}_{0\leq s\leq t}$ which are supported on S_{x_0}. This can be done be acting analogously to what we have done in Step 1 of the proof of Theorem 5.1.5; in particular we may define $\nu_t := \mathcal{Q}^*_{0,t}\delta_{x_0}$.[24] Note that $\nu_t(S_{x_0}) = 1$; indeed by Note 5.3.4 (second bullet point) we have that $\mathcal{Z}_t \in S_{x_0}$ almost surely when $\mathcal{Z}_0 = x_0$; hence

$$\nu_t(S_{x_0}) = \mathcal{Q}^*_{0,t}\delta_{x_0}(S_{x_0}) = \mathcal{Q}_{0,t}\mathbb{1}_{S_{x_0}}(x_0) = \mathbb{P}(\mathcal{Z}_t \in S_{x_0} | \mathcal{Z}_0 = x_0) = 1, \quad \text{for every } t \geq 0.$$

Moreover, analogously to Step 2 in the proof of Theorem 5.1.5, since the family $\{\nu_t\}_t$ is tight, there exists a diverging sequence $\{t_\ell\}_\ell$ such that ν_{t_ℓ} converges weakly to some probability measure μ_0 as t_ℓ tends to ∞.

- *Step 2*: By construction, the measure μ_0 is a measure on $\overline{S}_{x_0}$; we then consider the probability measure $\mu_0 \circ (W^\infty)^{-1}$.[25] The latter measure is supported on $\overline{S}_{\overline{x}}$. One needs to show that $\mu_0 \circ (W^\infty)^{-1} = \overline{\mu}^{S_{\overline{x}}}$. Recall that $\overline{\mu}^{S_{\overline{x}}}$ is the restriction of the measure $\mu^{S_{\overline{x}}}$ to $\overline{S}_{\overline{x}}$. The proof of this fact is deferred to Lemma A.2.12. Note

[24]Note that using the same argument we could define $\nu_t = \mathcal{Q}^*_{0,t}\delta_{x_0}$.
[25]Here $(W^\infty)^{-1}(A)$ denotes preimage of A.

that this is one of the places where we use that $x_0 \in \mathcal{I}_0(\overline{x})$. This implies that ν_t converges weakly to $\overline{\mu}^{S_{\overline{x}}} \circ W^\infty$ as t tends to ∞. Furthermore, by Hypothesis [**A.3**] we can take a sequence $\{t_\ell\}_\ell$ such that $t_\ell \nearrow \infty$ and $p^{x_0}_{t_\ell}$ converges weakly to some probability measure ν^{x_0}.

- *Step 3*: We show that ν^{x_0} is supported on $\overline{S}_{\overline{x}}$ and, when we restrict it to $\overline{S}_{\overline{x}}$, we have $\nu^{x_0}|_{\overline{S}_{\overline{x}}} = \mu_0 \circ (W^\infty)^{-1}$. Lemma A.2.13 is devoted to proving this fact.

Therefore, by Step 2 and the definition of $\mu^{S_{\overline{x}}}$ we have that

$$\nu^{x_0} = \mu^{S_{\overline{x}}}.$$

This implies that $p^{x_0}_t$ converges weakly to $\overline{\mu}^{S_{\overline{x}}}$ as t tends to ∞ for any $x \in \mathscr{S}_{x_0}$, that is, for every $f \in C_b(\mathbb{R}^N)$, $\mathcal{P}_t f(x_0)$ converges to $\mu^{S_{\overline{x}}}(f)$ as t tends to ∞. □

We now give a one dimensional example which satisfies all the assumptions we have made in this section. In particular, this example fits our framework in a non-trivial way as it exhibits many invariant measures.

Example 5.3.8. Consider the SDE

$$dZ^z_t = \sin(Z^z_t)dt + \sqrt{2}(1-\cos(Z^z_t)) \circ dB_t, \quad Z_0 = z, \quad Z_t \in \mathbb{R},$$

where $(B_t)_{t\geq 0}$ is a one-dimensional Wiener process. In this case $V_0 = \sin(z)\partial_z$, $V_1 = (1-\cos(z))\partial_z$ and we have

$$[V_1, V_0] = [(1-\cos(z))\partial_z, \sin(z)\partial_z] = \cos(z)(1-\cos(z))\partial_z - \sin(z)^2\partial_z = -V_1.$$

Therefore the vector fields V_0, V_1 satisfy the UFG condition; the above also shows that the obtuse angle condition (2.8) is satisfied, with $\lambda_0 = 1$. Moreover, it is easy to show that the function $(V_1\mathcal{P}_t f)(x)$ decays exponentially fast in time, i.e. λ_0 is big enough that (2.8) implies an estimate of the type (2.9) for the fields V_1. Because the coefficients of the equation are bounded the estimate is uniform on the whole real line, see [32, Proposition 3.1, Proposition 3.4 and Theorem 4.2] alternatively by a direct calculation, see [32, Example 4.4]. Since V_0 and V_1 both vanish whenever $z \in 2\pi\mathbb{Z}$ we have that the point measures $\delta_{2n\pi}$ are invariant measures for any $n \in \mathbb{Z}$. However there also exist invariant measures supported on $(2n\pi, 2(n+1)\pi)$ for any

$n \in \mathbb{Z}$. Indeed let

$$\rho_n(z) := \frac{\exp\left(-\frac{1}{1-\cos(z)}\right)}{C(1-\cos(z))}\mathbb{1}_{(2n\pi,2(n+1)\pi)}(z)$$

where C is the normalization constant and $\mathbb{1}_{(2n\pi,2(n+1)\pi)}(z)$ is the characteristic function of the interval $[2n\pi, 2(n+1)\pi)]$. By direct calculation one can verify that, for every $n \in \mathbb{Z}$, $\rho_n(z)$ satisfies the stationary Fokker-Planck equation $\mathcal{L}^*\rho_n = 0$, where

$$\mathcal{L}^*\rho_n(z) = -\partial_z(\sin(z)\rho_n(z)) + \partial_z\left[(1-\cos(z))\partial_z\left((1-\cos(z))\rho_n(z)\right)\right] .$$

Notice that if $X_0 \in [2n\pi, 2(n+1)\pi]$ (for some fixed $n \in \mathbb{Z}$) then $X_t \in [2n\pi, 2(n+1)\pi]$ for every $t \geq 0$. However, even if we restrict to one of the intervals $[2n\pi, 2(n+1)\pi]$, the process still admits three invariant measures on each one of such intervals. $\square$

Example 5.3.9 (Example 5.1.10 continued)**.** Recall that in this example $V_0 = \sin\zeta\partial_\zeta - kz\partial_z$ and $V_1 = \zeta\partial_z$. While V_0 is smooth, $V_0^{(\perp)}$ is not continuous. Indeed, for $\zeta \neq 0$ $V_0^{(\perp)}(z,\zeta) = -\sin(\zeta)\partial_\zeta$, however for $\zeta = 0$ $V_0^{(\perp)}(z,0) = V_0(z,0) = -kz\partial_z$. $\square$

We conclude this section by stating and proving Theorem 5.3.10 below. In order to state it, let us define the following equivalence relation on $\mathbb{R}^N$:

$$x \sim y \quad \Leftrightarrow \quad x \in S_y\,.$$

As customary, we denote by $[x]$ the equivalence class of x under the equivalence relation $\sim$. Note that by Lemma 2.3.8, if $x \sim y$ then also $e^{tV_0^{(\perp)}}x \sim e^{tV_0^{(\perp)}}y$, therefore the flow map

$$[x] \longrightarrow [e^{tV_0^{(\perp)}}x] =: e^{tV_0^{(\perp)}}[x] \tag{5.25}$$

is well defined. Let now q be the map $q : \mathbb{R}^N \to \mathbb{R}^N/\sim$, defined as $q(x) = [x]$. If we endow the quotient set $\mathbb{R}^N/\sim$ with the σ-algebra

$$\{E \subseteq \mathbb{R}^N/\sim \ \text{ such that } \ q^{-1}(E) \text{ is a Borel set of } \mathbb{R}^N\},$$

then q is a measurable map. If μ is a probability measure on $\mathbb{R}^N$, we define the pullback measure $\tilde{\mu}$ on $\mathbb{R}^N/\sim$ as $\tilde{\mu}(E) = \mu(q^{-1}(E))$ for all $E \subseteq \mathbb{R}^N/\sim$.

Theorem 5.3.10. *Consider the SDE* (1.1) *and the associated semigroup $\mathcal{P}_t$ and assume that the vector fields $V_0, \ldots, V_d$ satisfy the UFG condition. If μ is an invariant measure for $\mathcal{P}_t$, then $\tilde{\mu}$ is an invariant measure for the flow map* (5.25).

Proof of Theorem 5.3.10. Denote by $B_b(\mathbb{R}^N/\sim;\mathbb{R})$ to be the set of all bounded and measurable functions $f : \mathbb{R}^N/\sim \to \mathbb{R}$. If $f \in B_b(\mathbb{R}^N/\sim;\mathbb{R})$, then $f \circ q \in B_b(\mathbb{R}^N;\mathbb{R}^N)$, i.e. $f \circ q$ is a bounded and measurable function mapping from $\mathbb{R}^N$ to $\mathbb{R}^N$. By the definition of invariant measure, we have

$$\begin{aligned}
\int_{\mathbb{R}^N/\sim} f([x])\tilde{\mu}(d[x]) &= \int_{\mathbb{R}^N} f(q(x))\mu(dx) \\
&= \int_{\mathbb{R}^N} \left(\mathcal{P}_t(f \circ q)\right)(x)\mu(dx) \\
&= \int_{\mathbb{R}^N} \mathbb{E}_x[f(q(X_t^x))]\mu(dx)\,.
\end{aligned}$$

Let us now look more closely at the expected value on the right hand side of the above: for any bounded and measurable function h we can write

$$\begin{aligned}
\int_{\mathbb{R}^N} \mathbb{E}_x[h(X_t^x)]\mu(dx) &= \int_{\mathbb{R}^N} \mu(dx) \int_{\mathbb{R}^N} h(y)\mathbb{P}_x(X_t^x \in dy) \\
&= \int_{\mathbb{R}^N} \mu(dx) \int_{\overline{S_{e^{tV_0^{(\perp)}}x}}} h(y)\mathbb{P}_x(X_t^x \in dy) \\
&= \int_{\mathbb{R}^N} \mu(dx) \int_{S_{e^{tV_0^{(\perp)}}x}} h(y)\mathbb{P}_x(X_t^x \in dy) \\
&+ \int_{\mathbb{R}^N} \mu(dx) \int_{\partial S_{e^{tV_0^{(\perp)}}x}} h(y)\mathbb{P}_x(X_t^x \in dy)\,,
\end{aligned}$$

where the second equality follows from Proposition 3.2.3. Now the second term in the above vanishes by Proposition 3.2.7 (easy to prove for positive h, if h is not positive just split into positive and negative part). Indeed if $X_t^x \in \partial S_{e^{tV_0^{(\perp)}}x}$ then $X_t^x \in \partial \mathscr{S}_x$ (see Lemma 3.1.16 for a proof of this fact). Putting everything together we can write

$$\begin{aligned}\int_{\mathbb{R}^N/\sim} f([x])\tilde{\mu}(d[x]) &= \int_{\mathbb{R}^N} \mathbb{E}_x[f(e^{tV_0^{(\perp)}}([x]))]\mu(dx)\\ &= \int_{\mathbb{R}^N} f(e^{tV_0^{(\perp)}}(x))\mu(dx)\\ &= \int_{\mathbb{R}^N/\sim} f(e^{tV_0^{(\perp)}}([x]))\tilde{\mu}(d[x]),\end{aligned}$$

where the penultimate equality follows from the fact that the object on the second line is completely deterministic and the last equality holds by the definition of the measure $\tilde{\mu}$. This concludes the proof. □

Chapter 6

Existence of a density

Analogously to what we did for the study of the long-time behaviour, we split this section into two subsections. That is, in Section 6.1 we consider the setting of Section 5.1 and study SDEs of the form (5.1)-(5.3). In Section 6.2 we consider the general UFG-case. This section makes use of several notions from Malliavin calculus, we will recall only some basic facts and refer the reader to [35] for more detailed background material.

Let $\mathbb{D}^{k,p} \subseteq L^p(\Omega)$ denote the Malliavin Sobolev space, that is the domain of the kth order Malliavin derivative in the space $L^p(\Omega)$. We also define the space

$$\mathbb{D} = \bigcap_{p>1, k\in\mathbb{N}} \mathbb{D}^{k,p}.$$

We shall denote by $\mathbb{D}'$ the dual space of $\mathbb{D}$, that is the space of all continuous linear maps from $\mathbb{D}$ to $\mathbb{R}$. Let us recall the following lemma, which is quoted from [35, Theorem 2.2.1].

Lemma 6.0.1. *Fix $T > 0$, let $\{X_t\}_{t\in[0,T]}$ denote the solution of the SDE* (1.1) *and assume that $V_0, V_1, \dots, V_d$ are smooth vector fields which are globally Lipschitz. Then X_t^i belongs to $\mathbb{D}^{1,p}$ for any $t \in [0,T], p \geq 1$ and $i = 1, \dots, N$. Moreover, for all $0 \leq t \leq T, p \geq 1$*

$$\sup_{0\leq r\leq t} \mathbb{E}\left[\sup_{r\leq s\leq T} |D_r^j X_s^i|^p\right] < \infty$$

and the Malliavin derivative $D_r^j X_t^i$ satisfies the following SDE,

$$D_r^j X_t^i = V_j^i(X_r)+\sum_{k=1}^{N}\int_r^t \partial_{x^k} V_0^i(X_s) D_r^j(X_s^k)ds+\sqrt{2}\sum_{\ell=1}^{d}\sum_{k=1}^{N}\int_r^t \partial_{x^k} V_\ell^i(X_s) D_r^j(X_s^k)\circ dW_s^\ell, \tag{6.1}$$

for every $r \le t$.

Here we use the notation D^k to denote the Malliavin derivative operator with respect to the Brownian motion B^k.[26] Define the Malliavin matrix $\mathscr{M}_t = (\mathscr{M}_t^{ij})_{i,j=1}^N$ to be

$$\mathscr{M}_t^{ij} = \sum_{k=1}^{d}\int_0^t D_s^k(X_t^i) D_s^k(X_t^j)ds.$$

Again by [35, Section 2.3] we can rewrite the Malliavin matrix in terms of the Jacobian matrix $J_t := \frac{\partial X_t}{\partial x_0}$, details can be found in [35, Section 2.3]. There it is also shown that J_t is an invertible matrix and that the following holds

$$\mathscr{M}_t = J_t\left(\sum_{k=1}^{d}\int_0^t J_s^{-1}V_k(X_s)V_k(X_s)^T(J_s^{-1})^T ds\right)J_t^T = J_t\mathscr{C}_t J_t^T$$

where the matrix $\mathscr{C}_t$ is the reduced Malliavin covariance matrix defined as

$$\mathscr{C}_t = \sum_{k=1}^{d}\int_0^t J_s^{-1}V_k(X_s)V_k(X_s)^T(J_s^{-1})^T ds.$$

6.1 Existence of a density on a suitable hyperplane

In this section we consider the SDE (5.1)-(5.3). We shall also assume Hypothesis 5.1.1 [**H.1**], which states that the set of vector fields $\{V_{[\alpha]}(z_0,\zeta_0) : \alpha \in \mathcal{A}_m\}$ span the n-dimensional hyperplane $\mathcal{H}_{\zeta_0} := \{x = (z,\zeta) : \zeta = \zeta_0\}$ for all $(z_0,\zeta_0) \in \mathbb{R}^n \times \mathbb{R}$. In this setting it is clear that the law of $X_t = (Z_t,\zeta_t)$ does not admit a density with respect to Lebesgue measure on $\mathbb{R}^{n+1}$; indeed for each fixed t, ζ_t is a deterministic point which implies that $\mathbb{P}_x(X_t \in \mathbb{R}^n \times \{\zeta_t\}) = 1$ while $\mathbb{R}^n \times \{\zeta_t\}$ is a null set with respect to Lebesgue measure on $\mathbb{R}^{n+1}$. We prove that for every fixed $t \ge 0$ the law

[26] Note that D^k denotes the 1st order Malliavin derivative with respect to the kth Brownian motion and is not to be confused with the kth-order Malliavin derivative.

of the random variable Z_t admits a density with respect to Lebesgue measure on $\mathbb{R}^n$. In terms of the process X_t this implies that the law of X_t admits (for every fixed $t \geq 0$) a density with respect to the Lebesgue measure on the hyperplane $\mathcal{H}_{\zeta_t} := \{x = (z, \zeta) : \zeta = \zeta_t\}$. Moreover, since from Section 5.1 $X_t \in \mathcal{H}_{\zeta_t}$ almost surely, we have that $\mathcal{H}_{\zeta_t}$ is the maximal manifold such that X_t admits a density with respect to the volume element on such a manifold.

To prove that the law of Z_t admits a density we shall follow the same strategy of [35, Section 2.3]. Note that by Hypothesis 2.4.1 and Lemma 6.0.1 for each $t \geq 0$ and $i \in \{1, \ldots, n\}$ we have that Z_t^i and ζ_t belong to $\mathbb{D}^{1,p}$ for all $p \geq 1$. First we note that the solution $X_t = (Z_t, \zeta_t)$ admits a Malliavin derivative.

Lemma 6.1.1. *Let $\mathscr{M}_t$ denote the Malliavin matrix corresponding to the solution $X_t = (Z_t, \zeta_t)$ of the SDE* (5.1) - (5.3)*. Then $\mathscr{M}_t$ has the form*

$$\mathscr{M}_t = \begin{pmatrix} M_t & 0 \\ 0 & 0 \end{pmatrix} \tag{6.2}$$

where the matrix M_t is the Malliavin matrix corresponding to Z_t.

Proof of Lemma 6.1.1. The proof is deferred to Appendix A.2.5. □

In [35] it is shown that if the Malliavin matrix is invertible then the law of X_t admits a density on $\mathbb{R}^{n+1}$. We can see from (6.2) that the matrix is not invertible; however we show that the Malliavin matrix M_t corresponding to Z_t is invertible almost surely and hence the law of Z_t admits a density on $\mathbb{R}^n$, for every fixed $t > 0$.

Proposition 6.1.2. *The reduced Malliavin covariance matrix $\mathscr{C}_t$ corresponding to the solution $X_t = (Z_t, \zeta_t)$ of the SDE* (5.1) - (5.3) *is of the form*

$$\mathscr{C}_t = \begin{pmatrix} C_t & 0 \\ 0 & 0 \end{pmatrix},$$

where C_t is a random $n \times n$ symmetric matrix. Moreover, if we assume Hypothesis 5.1.1 **[H.1]** *holds then C_t is invertible $\mathbb{P}$-almost surely.*

Proof of Proposition 6.1.2. The proof is deferred to Appendix A.2.5. □

Theorem 6.1.3. *Assume Hypothesis 5.1.1* [**H.1**] *and let* $\{Z_t\}_{t\geq 0}$ *be the solution of* (5.1)*. Then the law of* Z_t *is absolutely continuous with respect to the Lebesgue measure on* $\mathbb{R}^n$*.*

Proof of Theorem 6.1.3. Note the Malliavin matrix corresponding to Z_t is M_t which is invertible, indeed $M_t = J_t C_t J_t^T$ and C_t is invertible by Proposition 6.1.2 therefore M_t is invertible since the product of invertible matrices is invertible. By [35, Theorem 2.1.2] we have that the law of Z_t is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^n$, for each $t > 0$. □

6.2 Existence of a density on integral submanifolds

We now return to studying the general UFG-case. As in the previous section we cannot expect that the law of X_t will in general admit a density with respect to Lesbegue measure on $\mathbb{R}^N$ and we will instead show that the law of X_t admits a density with respect to the volume element on a suitable manifold. Indeed, we shall show that the law of X_t admits a density with respect to the volume element on $S_{e^{tV_0^{(\perp)}}(x_0)}$. Note that by Proposition 3.2.3 we have $X_t^{(x)} \in \overline{S_{e^{tV_0^{(\perp)}}(x)}}$ almost surely. In this section we shall assume Hypothesis 5.3.3 [**A.1**] and that X_t cannot hit the boundary of the integral manifold $S_{e^{tV_0^{(\perp)}}(x)}$, that is $X_t^{(x)} \in S_{e^{tV_0^{(\perp)}}(x)}$, almost surely. In the first and second comment in Note 5.3.4 it is shown that under Hypothesis 5.3.3 [**A.4**] implies that X_t cannot hit the boundary of the maximal integral submanifold.

Recall from Section 5.2 the process $\{\mathcal{Z}_t\}_t$ defined by (5.17). Since $e^{-tV_0^{(\perp)}}$ is a diffeomorphism the law of X_t admits a density with respect to the volume element on $S_{e^{tV_0^{(\perp)}}(x_0)}$ if and only if the law of $\mathcal{Z}_t$ admits a density with respect to the volume element on S_{x_0}. Let $\mathcal{V}_{[\alpha],t}$ be defined as in (5.21), then recall that the process $\{\mathcal{Z}_t\}_{t\geq 0}$ satisfies the SDE (5.19). Now we wish to apply [42, Theorem 3.4] to show that the law of $\{\mathcal{Z}_t\}_{t\geq 0}$ admits a density with respect to the volume measure on S_{x_0}. However, as noted in [38], there is a mistake in the proof of [42, Theorem 3.4], in particular the form of the Hörmander condition given by [42, Assumption (H)] is not sufficient for the conclusions of [42, Theorem 3.4] to hold. More precisely, they rely upon [64, Theorem 1.1.3] to show that [42, Assumption (H)] implies a suitable integration by parts formula, which is shown to be incorrect by [38]. However under our conditions there is an integration by parts formula as shown in [50, Section 3]. Therefore we may use the strategy given in [42] and the results of [50] to prove that the law of $\mathcal{Z}_t$ admits a density with respect to the volume measure on S_{x_0}.

A vital tool for this argument is the integration by parts formula proved in [50, Theorem 3.10]; namely, for $\Phi \in \mathbb{D}$, $\alpha_1, \ldots, \alpha_M \in \mathcal{A}_m$ and $f \in C_V^\infty(\mathbb{R}^N)$ we have

$$\mathbb{E}_x \left[\Phi V_{[\alpha_1]} \ldots V_{[\alpha_M]} f(X_t)\right] = t^{\frac{-\|\alpha_1\| - \ldots - \|\alpha_M\|}{2}} \mathbb{E}_x[\Phi_{\alpha_1,\ldots,\alpha_M}(t,x) f(X_t)],$$

for some random variable $\Phi_{\alpha_1,\ldots,\alpha_M}(t,x)$ independent of f. By taking $f = g \circ e^{-tV_0^{(\perp)}}$ we have

$$\mathbb{E}_x \left[\Phi \mathcal{V}_{[\alpha_1],t} \ldots \mathcal{V}_{[\alpha_M],t} g(\mathcal{Z}_t)\right] = t^{\frac{-\|\alpha_1\| - \ldots - \|\alpha_M\|}{2}} \mathbb{E}_x[\Phi_{\alpha_1,\ldots,\alpha_M}(t,x) g(\mathcal{Z}_t)], \tag{6.3}$$

for any $g \in C_V^\infty(\mathbb{R}^N)$.

Let us denote by $\mathcal{E}(S_{x_0})$ the space of all distributions on S_{x_0} with compact support. Recall we can consider a smooth function f to be a distribution F_f by setting

$$\langle F_f, \phi \rangle = \int_S f(x)\phi(x)\lambda_{S_{x_0}}(dx), \qquad \text{for any } \phi \in C_c^\infty(S_{x_0})$$

where $\lambda_{S_{x_0}}$ denotes the volume measure on S_{x_0}.

Lemma 6.2.1. *Assume that $\mathcal{Z}_t$ satisfies (6.3). Then there exists a map $\Psi_t : \mathcal{E}(S_{x_0}) \to \mathbb{D}'$ with the following properties*

1. *If $f \in C_c^\infty(S_{x_0})$ then $\Psi_t(f) = f(\mathcal{Z}_t)$. Note that $f(\mathcal{Z}_t)$ is identified as an element in $\mathbb{D}'$ by setting $\langle f(\mathcal{Z}_t), G\rangle = \mathbb{E}[f(\mathcal{Z}_t)G]$ for any $G \in \mathbb{D}$.*

2. *The map Ψ_t is continuous as a map from $\mathcal{E}(S_{x_0})$ to $\mathbb{D}'$.*

For a proof see [65, Proposition 2.1].

Now we shall state some properties of the map Ψ, as proven in [66, Proposition 2].

Proposition 6.2.2. *Fix $t > 0$ and let $\mathcal{Z}_t$ be such that the map Ψ_t is well defined for every distribution f. Then let I be some open set*

1. *If $I \ni s \mapsto F_s$ is continuous (continuously differentiable), then $I \ni s \mapsto \Psi_t(F_s)$ is continuous (resp. continuously differentiable). In particular, for every $G \in \mathbb{D}$ the map $I \ni s \mapsto \langle \Psi_t(F_s), G\rangle$ is continuous and respectively continuously differentiable and*

$$\left\langle \Psi_t\left(\frac{dF_s}{ds}\right), G\right\rangle = \frac{d}{ds}\left\langle \Psi_t(F_s), G\right\rangle .$$

2. *If $I \ni s \mapsto F_s$ is continuous then for every $G \in \mathbb{D}$*

$$\left\langle \Psi_t\left(\int_I F_s ds\right), G\right\rangle = \int_I \langle \Psi_t(T_s), G\rangle ds$$

where $\int_I T_s ds$ is a tempered distribution and is defined by $\langle \int_I T_s ds, \phi\rangle = \int_I \langle T_s, \phi\rangle ds$.

We can show that the law of $\mathcal{Z}_t$ admits a density.

Proposition 6.2.3. *Assume Hypothesis 5.3.3* [**A.1**]*, and assume that $X_t^{(x_0)} \in S_{e^{tV_0^{(\perp)}}(x_0)}$ almost surely. Then for each $t > 0$ the law of $\mathcal{Z}_t^{(x_0)}$ admits a density with respect to the volume element on S_{x_0}.*

Proof. Note that the map $x \mapsto \delta_x$ is smooth, moreover its (weak) derivative $\frac{d}{x^i}\delta_x$ is given by $D_i\delta_x$, where $D_i\delta_x$ is defined by $\langle D_i\delta_x, \phi\rangle = -\partial_{x^i}\phi(x)$ for all ϕ. Therefore $\Psi_t(\delta_x)$ is smooth and in particular $p(x) := \langle \Psi_t(\delta_x), 1\rangle$ is smooth. It remains to show that $p(x)$ is the density of the law of X_t. Take $\phi \in C_c^\infty(S)$ then

$$\begin{aligned}\int_S \phi(x)p(x)\lambda_{S_{x_0}}(dx) &= \int_S \phi(x)\langle \Psi(\delta_x), 1\rangle \lambda_{S_{x_0}}(dx)\\ &= \left\langle \Psi_t\left(\int_{S_{x_0}} \phi(x)\delta_x\lambda_{S_{x_0}}(dx)\right), 1\right\rangle .\end{aligned}$$

Now for $f \in C_c^\infty(S_{x_0})$ we have

$$\left\langle \int_{S_{x_0}} \phi(x)\delta_x\lambda_{S_{x_0}}(dx), f\right\rangle = \int_{S_{x_0}} \phi(x)\langle \delta_x, f\rangle \lambda_{S_{x_0}}(dx) = \int_S \phi(x)f(x)\lambda_{S_{x_0}}(dx).$$

Therefore $\int_{S_{x_0}} \phi(x)\delta_x\lambda_{S_{x_0}}(dx) = F_\phi$, and in particular $\Psi_t\left(\int_{S_{x_0}} \phi(x)\delta_x\lambda_{S_{x_0}}(dx)\right) =$

$\phi(\mathcal{Z}_t)$. Now we have

$$\int_S \phi(x)p(x)\lambda_S(dx) = \langle \phi(\mathcal{Z}_t), 1\rangle = \mathbb{E}[\phi(\mathcal{Z}_t)].$$

$\square$

Theorem 6.2.4. *Assume the vector fields $V_0, V_1, \ldots, V_d$ are uniformly Lipschitz, satisfy the UFG condition and assume that $X_t^{(x_0)} \in S_{e^{tV_0^{(\perp)}}(x_0)}$ almost surely.*[27] *Then for each $t > 0$ the law of $X_t^{(x_0)}$ admits a density with respect to the volume element on $S_{e^{tV_0^{(\perp)}}(x_0)}$.*

[27]As we have already mentioned, the latter fact follows for example from assuming Hypothesis 5.3.3 [**A.4**].

Conclusions and future work

For this thesis we have been interested in problems related to SDEs which exhibit multiple invariant measures. We concentrated on SDEs of UFG type, which is more general than the Uniform Parabolic Hörmander Condition, and we have explored the behaviour of such SDEs. Namely, we have shown that there is a change of coordinates such that we can (locally) write the SDE (1.1) in the form "ODE+SDE" (3.7)-(3.9). In Section 5.1 we consider the case when the SDE can be globally expressed in the form "ODE+SDE", that is we consider the SDE (5.1)-(5.3). In such a situation we can fully describe the long time behaviour. Indeed, since the ODE component is one-dimensional and autonomous it must either converge to a point or diverge to $\pm\infty$. Then by Theorem 5.1.6 we see that the SDE can only support an invariant measure on the set $\{x = (z, \zeta) : W_0(\zeta) = 0\}$ and for each point $\overline{\zeta}$ with $W_0(\overline{\zeta}) = 0$ there is a unique invariant measure $\mu^{\overline{\zeta}}$ supported on the hyperplane $\mathcal{H}_{\overline{\zeta}}$ and the basin of attraction of $\mu^{\overline{\zeta}}$ is given by $\mathcal{I}_0(\overline{\zeta})$.

In Section 5.2 we extend these results to the more general situation where the SDE is only locally of the form "ODE+SDE". This analysis relies on the observation that the "ODE" component (3.8) corresponds to the globally defined curve $e^{tV_0^{(\perp)}}(x)$ where $V_0^{(\perp)}$ is defined by (1.12). In contrast to the situation considered in Section 5.1, in the more general situation there are many behaviours which the ODE can exhibit, for example convergence, periodicity, limit cycles, or chaotic behaviour. In this thesis, we addressed the case when is ODE convergent. A natural future direction is to investigate each of the different cases listed which we expect to lead to understanding a rich structure which incorporates a wide variety of long time behaviours. Further, another natural question to address would be to determine the rate of convergence of the SDE to equilibria. Indeed, from the proof of Lemma 3.2.9 we see that $\mathcal{P}_t f(x) - \mathcal{P}_t f(y)$ converges to zero exponentially fast for $x, y \in S$ under the assumptions of Lemma 3.2.9. However the constant appearing on the right hand side of (A.18) depends on the length of the shortest path between x and y and in general we can not control this distance. Therefore, we may expect that, under the UFG condition and assuming (2.9) holds, the semigroup may converge to equilibria exponentially fast however to prove this will require further work.

Another interesting extension of this work would be to study the case where the vector fields are not necessarily smooth. It has been shown in [50] that if the vector fields are only k times continuously differentiable then some smoothing properties still hold. However in the work of Rampazzo and Sussmann, see [67], they develop a concept for commutators of vector fields which are only Lipschitz, and show that some geometrical concepts still hold. Investigating how we can interpret these geometrical results for SDEs with only Lipschitz coefficients, one could expect to be able to determine what conditions we require to understand the asymptotic behaviour of the process.

In finite dimensions we found investigating systems which are defined globally as (1.13)-(1.14) gave an insight into the more general case in which the representation is only local. It would be of interest to investigate the corresponding infinite dimensional scenario.

In Chapter 4 we introduced a pathwise approach to obtaining long time derivative estimates of the form (1.15). However, for this chapter, we only considered the case when $N = d = 1$ which enables us to significantly simplify the argument. We are currently working on extending these results to the higher dimensional setting, this will be the object of [43].

Appendix A

Appendix

A.1 Some technical results

We gather in this appendix some auxiliary results. In particular, Appendix A.1.1 recalls why the PHC is important in a probabilistic context. Appendix A.1.2 shows the implications between the UHC, PHC and UFG condition. Appendix A.1.3 contains background material about the topology of the orbits of finitely generated smooth distributions. Appendix A.1.4 reports some known smoothing results on UFG semigroups, which are often used in the proofs of Appendix A.2. Appendix A.1.5 contains precise statements and proofs of further technical facts which would have been cumbersome (and detracting from the main line of thought) if presented in the main body of the work.

A.1.1 Probabilistic implications of the parabolic Hörmander condition

In this section we summarise the probabilistic implications of (**PHC**), this can be found in more detail in [34].

We start by recalling the definition of hypoellipticity. Note that throughout this section of the Appendix we use the word "distribution" in an analytic sense, i/i not in the geometric sense from the rest of the thesis.

Definition A.1.1. A linear differential operator $\mathcal{T}$ on $\mathcal{O}$ (for some open set $\mathcal{O} \subseteq \mathbb{R}^N$) with C^∞ coefficients is *hypoelliptic* if sing supp u = sing supp $\mathcal{T}u$ for every distribution u on $\mathcal{O}$. The singular support of u, denoted by sing supp u, is the set of points in $\mathcal{O}$ having no open neighbourhood such that the restriction of u to such a neighbourhood is a C^∞ function. In other words $\mathcal{T}$ is hypoelliptic if the distribution u is a smooth function on any open set where $\mathcal{T}u$ is a smooth function.

One of the reasons why the concept of hypoellipticity is of great importance is because it allows to prove the existence of a smooth density for the law of the process (1.1). To show this fact, let $f(t,x) \in C_c^\infty(\mathbb{R} \times \mathbb{R}^N)$ with $f(0,x) = 0$ for all $x \in \mathbb{R}^N$. By Itô's formula

$$f(t, X_t) - f(0, X_0) = \int_0^t \left(\frac{\partial}{\partial s} + \mathcal{L} \right) f(s, X_s) ds + \sum_{i=1}^d \int_0^t V_i f(s, X_s) dW_s.$$

So if t is big enough $f(t, X_t) = f(0, X_0) = 0$ and taking expectation on both sides of the above equality we get

$$0 = \mathbb{E} \int_0^t \left(\frac{\partial}{\partial s} + \mathcal{L} \right) f(s, X_s) ds = \int_0^t ds \int_{\mathbb{R}^N} \left[\left(\frac{\partial}{\partial s} + \mathcal{L} \right) f(s, y) \right] \mathbb{P}(X_s \in dy). \tag{A.1}$$

The above is justified since the vector field $V_0, V_1, \ldots, V_d$ are smooth. Formally integrating by parts gives

$$\int_0^t ds \int_{\mathbb{R}^N} \left[\left(-\frac{\partial}{\partial s} + \mathcal{L}^* \right) \mathbb{P}(X_s \in dy) \right] f(s, y) = 0 \quad \forall f \in C_c^\infty((0, \infty) \times \mathbb{R}^N), \tag{A.2}$$

where $\mathcal{L}^*$ is the formal L^2-adjoint of $\mathcal{L}$. More rigorously, the measure $\mathbb{P}(X_s \in dy)ds$ (seen as a measure on $(0,\infty) \times \mathbb{R}^N$, which is finite on compacts[28]) induces a distribution on $(0,\infty) \times \mathbb{R}^n$, which we denote by $u_{\mathbb{P}}$. The right hand side of (A.1) is just the action of $u_{\mathbb{P}}$ on the test function $\left(\frac{\partial}{\partial s} + \mathcal{L} \right) f(s, y)$. Therefore (A.1) is equivalent to

$$u_{\mathbb{P}} \left(\left(\frac{\partial}{\partial s} + \mathcal{L} \right) f \right) = 0, \qquad \forall f \in C_c^\infty((0, \infty) \times \mathbb{R}^N).$$

[28]For each fixed $s > 0$, $\mathbb{P}(X_s \in dy)ds$ is a probability measure on $\mathbb{R}^n$. However, $\int_a^b ds \int_{\mathbb{R}^N} \mathbb{P}(X_s \in dy) = b - a$, which is finite for any $b > a > 0$.

By definition, this means

$$\left[\left(-\frac{\partial}{\partial s}+\mathcal{L}^*\right)u_{\mathbb{P}}\right](f)=0, \qquad \forall f\in C_c^\infty((0,\infty)\times\mathbb{R}^N).$$

That is, the distribution $\left(-\frac{\partial}{\partial s}+\mathcal{L}^*\right)u_{\mathbb{P}}$ is the null distribution:

$$\left(-\frac{\partial}{\partial s}+\mathcal{L}^*\right)u_{\mathbb{P}}=0.$$

Therefore, if $\partial_t-\mathcal{L}^*$ is hypoelliptic, then $u_{\mathbb{P}}$ is a C^∞ function; hence $\mathbb{P}(X_t\in dy)$ has a density, $p_t(y)$, which is a C^∞ function. In particular, from the above equality, $p_t(y)$ satisfies the Fokker Planck equation.

In his seminal paper [3], Hörmander gave a sufficient condition for a second order differential operator of the form (1.4) to be hypoelliptic.

Theorem A.1.2 (Hörmander's theorem)**.** *Let $V_0,V_1,\ldots,V_d$ be smooth vector fields on $\mathbb{R}^N$ and consider the second order differential operator $\mathcal{L}$ of the form* (1.4)*. If the Lie algebra generated by $\{V_0,V_1,...,V_d\}$, Lie$\{V_i\}_{0\leq i\leq d}$, is full for every $x\in\mathbb{R}^N$ then $\mathcal{L}$ is hypoelliptic.*

By Hörmander's Theorem, if

$$\mathrm{Lie}\{\partial_t+V_0,V_1,...,V_d\},$$

is full on $\mathbb{R}^N\times(0,\infty)$, then $\partial_t-\mathcal{L}^*$ is hypoelliptic on $\mathbb{R}^N\times(0,\infty)$. To check the hypoellipticity of such an operator, the following lemma is quite useful.

Lemma A.1.3. Lie$\{\partial_t+V_0,V_1,...,V_d\}$ *is full at each point of* $\mathbb{R}^N\times(0,\infty)$ *if and only if*

$$\mathrm{span}\{V_{[\alpha]}(x):\alpha\in\mathcal{A}\} \text{ is full at each point } x\in\mathbb{R}^N. \tag{A.3}$$

Recall Condition (A.3) is the PHC. In conclusion, one has the following.

Theorem A.1.4 (Smoothness of the density)**.** *Assume that X_t is the solution to* (1.1) *that* (**PHC**) *holds. Then the random vector X_t has an infinitely differentiable density for all $t>0$.*

The above Theorem A.1.4 can be found in [35, Theorem 2.3.2].

Note A.1.5. It should be clear from what we have just said but we would like to emphasize that the hypoellipticity of $\mathcal{L}^*$ on $\mathbb{R}^N$ is not the same thing as the hypoellipticity of $\partial_t - \mathcal{L}^*$ on $(0, \infty) \times \mathbb{R}^N$ and only the latter implies existence of a density. We illustrate this fact with the following simple example. Consider the equation in $\mathbb{R}$

$$dX_t = dt, \quad X_0 = x_0.$$

Then $\mathcal{L} = \partial_y =: V_0$, $\mathcal{L}^* = -\partial_y = -V_0$. Clearly $\mathcal{L}^*$ is hypoelliptic on $\mathbb{R}$. However the solution $X_t = t + x_0$ does not have a density; indeed, $\partial_t - \mathcal{L}^*$ is not hypoelliptic on $(0, \infty) \times \mathbb{R}$. Consistently, observe that the Lie algebra appearing in (**PHC**) is, in this case, simply equal to $\{0\}$ at every point of $\mathbb{R}$. To show that $\partial_t - \mathcal{L}^*$ is not hypoelliptic, notice that $\partial_t - \mathcal{L}^* = \partial_t + \partial_y$. Consider for example the solution of the Cauchy problem $(\partial_t + \partial_y)u = 0$, $u(0, y) = f(y)$. The solution of this Cauchy problem is $u(t, y) = f(y - t)$. If f is not C^∞, then u is not C^∞ either. □

Finally, another important result.

Proposition A.1.6. *Let $\mathcal{L}$ be the generator of the SDE* (1.1) *and suppose* (**PHC**) *holds. Then $\mathcal{L}, \mathcal{L}^*, \partial_t - \mathcal{L}, \partial_t - \mathcal{L}^*$ are hypoelliptic. The transition probabilities $p_t(x, dy)$ of the process have C^∞ densities, $p_t(x, y)$ (and the density is C^∞ in t, x and y). Furthermore, the semigroup associated to the Markov process X_t is strong Feller. The invariant measures, if they exist, have C^∞ densities as well.*

The above can be found in [4, Corollary 7.2].

A.1.2 Parabolic Hörmander's condition and UFG condition

This appendix is devoted to showing the validity of the implications (2.6) and (2.7). We begin with a simple preliminary lemma.

Lemma A.1.7. *If a collection $\{V_i : i = 0, ..., d\}$ of C^∞ vector fields satisfy the UFG condition* (**UFG**) *for some $m \in \mathbb{N}$ then for any $n \geq m$ and $\alpha \in \mathcal{A}_n$ there exist $\phi_{\alpha,\beta} \in C_V^\infty$ such that*

$$V_{[\alpha]}(x) = \sum_{\beta \in \mathcal{A}_m} \phi_{\alpha,\beta}(x) V_{[\beta]}(x).$$

Proof of Lemma A.1.7. We prove this by induction, the result follows for $n = m$ by the (**UFG**) condition. Assume the result holds for some $n \geq m$. Then let $\alpha = \alpha' * i$

for $\alpha' \in \mathcal{A}_n, i \in \{0, 1, 2, ..., d\}$. By the inductive hypothesis we may write

$$V_{[\alpha']}(x) = \sum_{\beta \in \mathcal{A}_m} \phi_{\alpha',\beta}(x) V_{[\beta]}(x).$$

Since commutators are linear in the first argument we get

$$V_{[\alpha]} = [V_{[\alpha']}, V_i] = \sum_{\beta \in \mathcal{A}_m} [\phi_{\alpha',\beta} V_{[\beta]}, V_i].$$

Note by the definition of the commutator and since V_i is a first order differential operator

$$\begin{aligned}
[\phi_{\alpha',\beta} V_{[\beta]}, V_i] &= \phi_{\alpha',\beta} V_{[\beta]} V_i - V_i(\phi_{\alpha',\beta} V_{[\beta]}) \\
&= \phi_{\alpha',\beta} V_{[\beta]} V_i - V_i(\phi_{\alpha',\beta}) V_{[\beta]} - \phi_{\alpha',\beta} V_i(V_{[\beta]}) \\
&= \phi_{\alpha',\beta} [V_{[\beta]}, V_i] - V_i(\phi_{\alpha',\beta}) V_{[\beta]}.
\end{aligned}$$

As $\beta \in \mathcal{A}_n$ and $[V_{[\beta]}, V_i]$ can be expressed in the required form by the inductive hypothesis the result follows. □

It is immediate to see that the Parabolic Hörmander Condition (**PHC**) implies the Hörmander condition (**HC**) so we focus on the first implication in (2.6).

Lemma A.1.8 (Uniform Parabolic Hörmander implies parabolic Hörmander)**.** *Suppose a collection $\{V_i : i = 0, ..., d\}$ of vector fields satisfies* (**UPHC**) *then it also satisfies* (**PHC**).

Proof of Lemma A.1.8. Suppose for a contradiction that (**PHC**) does not hold, then there exists $x \in \mathbb{R}^N$ with

$$\bigcup_{j \geq 1} \operatorname{span}\{\mathfrak{L}_j(x)\} \neq \mathbb{R}^N.$$

Therefore, there exists $\xi \in \mathbb{R}^N$ which is orthogonal to $\bigcup_{j\geq 1} \operatorname{span}\{\mathfrak{L}_j(x)\}$ and has $|\xi| = 1$. By the definition of the set $\mathcal{L}_j(x)$ this gives $|V_{[\alpha]}(x) \cdot \xi|^2 = 0$ for all $\alpha \in \mathcal{A}_m$. In particular,

$$\sum_{\alpha \in \mathcal{A}_m} |V_{[\alpha]}(x) \cdot \xi|^2 = 0.$$

However this contradicts (**UPHC**), hence the parabolic Hörmander condition must be satisfied. □

The following example shows that the reverse implication is not true. The principle behind the example is that under the PHC for each fixed point we only need a finite number of commutators to generate enough vector fields to span $\mathbb{R}^N$, however as the point varies we may require an more and more commutators. Whereas, under the UPHC there must be a finite number of vector fields which span $\mathbb{R}^N$ for every point $x \in \mathbb{R}^N$.

Example A.1.9. Let $N = d = 1$ and consider the one-dimensional SDE

$$dX_t = dt + \sqrt{2} V_1(X_t) \circ dB_t.$$

Here $V_0 = \partial_x$. We shall construct the diffusion coefficient V_1 such that for each $n \in \mathbb{N}$ there is a neighbourhood $\mathcal{O}_n$ with

$$V_1(x) = \frac{1}{n!}(x-n)^n \quad \text{ for all } x \in \mathcal{O}_n.$$

We may assume that there are no points $x \in \mathbb{R} \setminus \mathbb{N}$ with $V_1(x) = V_1'(x) = 0$, in which case we only need to test (**PHC**) and (**UPHC**) at the points $n \in \mathbb{N}$.

Fix some $n \in \mathbb{N}$. Recall the $(k+1)$-tuple $\alpha_{1,k}$ defined in Example 2.2.7. Now

$$V_{[\alpha_{1,k}]}(n) = V_{[(1,\underbrace{0,\ldots,0}_{k \text{ times}})]} = V_1^{(k)}(n) = \delta_{k,n}$$

where $\delta_{k,n}$ is the Kronecker delta, i.e. $\delta_{k,n}$ equals one if $k = n$ and zero otherwise. Since $V_{[\alpha_{1,n}]}(n) = 1$, (**PHC**) holds. On the other hand, suppose (**UPHC**) is satisfied. Let ℓ be as in (**UPHC**) and set $n = \ell + 1$ then $V_{[\alpha]}(n) = 0$ for all $\alpha \in \mathcal{A}_\ell$. This gives a contradiction and hence (**UPHC**) is not satisfied. □

Lemma A.1.10 (Uniform Parabolic Hörmander Condition implies UFG)**.** *Suppose a collection $\{V_i : i = 0, ..., d\}$ of bounded vector fields satisfies the Uniform Parabolic Hörmander condition then it also satisfies the UFG condition.*

Proof of Lemma A.1.10. By Lemma A.1.8 we have that (**PHC**) is satisfied so span$\{V_\alpha :$

$\alpha \in \mathcal{A}_m\} = \mathbb{R}^N$. Hence for $\alpha \in \mathcal{A}_m$ there exist functions $\phi_{\alpha,\beta} : \mathbb{R}^N \to \mathbb{R}$ such that

$$V_{[\alpha]}(x) = \sum_{\beta \in \mathcal{A}_m} \phi_{\alpha,\beta}(x) V_{[\beta]}(x). \tag{A.4}$$

Define the $|\mathcal{A}_m| \times m$ matrix valued function $\boldsymbol{V}(x) = (V_{[\beta]}(x))_{\beta \in \mathcal{A}_m}$ and a $|\mathcal{A}_m|$-dimensional vector field $\boldsymbol{\phi}_\alpha(x) = (\phi_{\alpha,\beta}(x))_{\beta \in \mathcal{A}_m}$, then we may write (A.4) as

$$V_{[\alpha]} = \boldsymbol{V}^T(x)\boldsymbol{\phi}_\alpha(x)$$

Multiply this equation by $\boldsymbol{V}$ to get

$$\boldsymbol{V}(x)V_{[\alpha]}(x) = \boldsymbol{V}(x)\boldsymbol{V}^T(x)\boldsymbol{\phi}.$$

Note that (**UPHC**) can be expressed as

$$\inf\{\xi^T\boldsymbol{V}(x)\boldsymbol{V}^T(x)\xi : \xi, x \in \mathbb{R}^N, |\xi| = 1\} = \kappa > 0.$$

In particular, since $\boldsymbol{V}\boldsymbol{V}^T$ is symmetric the above gives that $\det(\boldsymbol{V}(x)\boldsymbol{V}^T(x)) \geq \kappa > 0$ for all $x \in \mathbb{R}^N$, hence $\boldsymbol{V}\boldsymbol{V}^T$ is invertible. Note that the inverse of a matrix of the form $\boldsymbol{V}\boldsymbol{V}^T$ is of the form $\frac{1}{\det(\boldsymbol{V}\boldsymbol{V}^T)}\boldsymbol{M}$ where $\boldsymbol{M}$ is a matrix with entries formed by polynomial terms of entries in $\boldsymbol{V}$; in particular since $\boldsymbol{V}$ has C_b^∞ terms so does $\boldsymbol{M}$. We also can see that $\det(\boldsymbol{V}\boldsymbol{V}^T)$ is C^∞ and $|\det(\boldsymbol{V}\boldsymbol{V}^T)| \geq c$ and hence $(\boldsymbol{V}\boldsymbol{V}^T)^{-1}$ has C_b^∞ coefficients. Now we may write $\boldsymbol{\phi}_\alpha$ as a product of matrices with C_b^∞ coefficients

$$(\boldsymbol{V}(x)\boldsymbol{V}^T(x))^{-1}\boldsymbol{V}(x)V_{[\alpha]}(x) = \boldsymbol{\phi}.$$

Therefore $\phi_{\alpha,\beta} \in C_b^\infty(\mathbb{R}^N)$ for all $\alpha \in \mathcal{A}, \beta \in \mathcal{A}_m$, and in particular (**UFG**) is satisfied. $\square$

To see that the converse fails consider the case where $N = 1, d = 0$ and $V_0 = \partial_x$ then the UFG condition is satisfied but neither (**UPHC**) or even (**PHC**) hold.

We summarise Lemma A.1.8 and Lemma A.1.10 in following Proposition.

Proposition A.1.11. *If the vector fields $V_{[\alpha]} \in C_b^\infty$ then the following implications* *hold.*

$$(\mathbf{UPHC}) \Rightarrow (\mathbf{PHC}) \Rightarrow (\mathbf{HC})$$

$$(\mathbf{UPHC}) \Rightarrow (\mathbf{UFG}).$$

A.1.3 Topology of orbits

Here we give a brief justification of the reason why we make the standing assumption [**SA.2**]. In short, assuming that the manifold topology of the manifolds $\mathscr{S}$ is the Euclidean topology is equivalent to assuming that such manifolds are embedded manifolds. In full generality, as explained in [58, page 78], elements of a global partition induced by distributions which enjoy the integral manifold property are immersed manifolds. We briefly explain the difference between an embedded and an immersed manifold. A detailed treatment of the matter can be found in [58, Appendix A.2 and Appendix A.4]. Let $F : \mathcal{M}_1 \to \mathbb{R}^N$ be a continuous mapping of topological spaces and let $\mathcal{M}_2 = F(\mathcal{M}_1)$. $\mathcal{M}_2$ can be endowed with two topologies: i) the topology of $\mathcal{M}_2$ as a subset of the Euclidean space $\mathbb{R}^N$, so that the open sets in this topology are the sets O of the form $O = O' \cap \mathcal{M}_2$ for some O' which is open in the Euclidean topology of $\mathbb{R}^N$; ii) the topology induced by $\mathcal{M}_1$, where the open sets are the sets U of the form $U = F(U')$, for some U' which is open in the topology of $\mathcal{M}_1$. In general, the latter topology is stronger than the former. With this premise, one can give the following definition.

Definition A.1.12. Let $F : \mathcal{M}_1 \to \mathbb{R}^N$ be a smooth mapping of manifolds. F is an *immersion* if it is injective and $rank(\mathcal{J}_p F) = dim(\mathcal{M}_1)$ for every $p \in \mathcal{M}_1$. F is an *embedding* if it is an immersion and the topology induced on $\mathcal{M}_2 = F(\mathcal{M}_1)$ by the one on $\mathcal{M}_1$ coincides with the Euclidean topology of $\mathcal{M}_2$ as a subset of $\mathbb{R}^N$.

The reason why we consider only the case in which the manifolds of the partition are embeddings comes mostly from the need to use the Stroock and Varadhan support theorem: the closure appearing in the statement of such a theorem is intended in Euclidean sense. If the manifold topology was not the Euclidean topology we would have to consider two closures, the closure in the Euclidean topology and the closure in the manifold topology. This would make the exposition much more cloudy. Moreover we point out that in all our examples the manifolds at hand are embedded manifolds. It is possible that, under the assumption of this paper that the vector $V_0^{(\perp)}$ is smooth and Lipshitz and that the integral curves of $V_0^{(\perp)}$ are convergent, one may prove that the orbits $\mathscr{S}$ are indeed embedded manifolds. But this is beyond the scope of this thesis.

A.1.4 Known facts about UFG semigroups

In this appendix we gather some known facts that we use frequently.

[F.1] A semigroup $\mathcal{P}_t$ of bounded operators is Markov if

$$\mathcal{P}_t 1 = 1 \qquad \text{and} \qquad \mathcal{P}_t f \geq 0 \text{ when } f \geq 0\,,$$

where, in the above, 1 denotes the constant function identically equal to one. Denoting by $\|\cdot\|_\infty$ the supremum norm, the above implies that if $\|f\|_\infty < \infty$ then $\|\mathcal{P}_t f\|_\infty \leq \|f\|_\infty$, i.e. the semigroup is a contraction in the supremum norm. Similarly the two parameter semigroups $\{Q_{s,t}\}_{0\leq s\leq t}$ and $\{\mathcal{Q}_{s,t}\}_{0\leq s\leq t}$, considered in Section 5.1 and Section 5.2, are both contractive in the supremum norm.

[F.2] Note that if the vector fields $V_0, V_1 \ldots, V_d$ satisfy the parabolic Hörmander condition then for any $f \in C_b(\mathbb{R}^N)$ then $\mathcal{P}_t f(x)$ is smooth in all directions in $\mathbb{R}^N$ and moreover is smooth in t. This is not generally the case if we assume UFG condition. However we have that for any $f \in C_b(\mathbb{R}^N)$ and $t > 0$ the function $x \mapsto \mathcal{P}_t f(x)$ is differentiable in the directions $V_{[\alpha]}$ for any $\alpha \in \mathcal{A}$. Moreover for any compact set K, $t > 0$ there exists $C(K) > 0, \omega > 0$ such that

$$\sup_{x\in K} \left|V_{[\alpha]}\mathcal{P}_t f(x)\right| \leq C(K) e^{\omega t} t^{-\|\alpha\|/2} \left|f\right|_\infty .$$

If the vector fields $V_{[\alpha]}$ are bounded then the above estimate holds uniformly on $\mathbb{R}^N$, for details see [50, Chapter 3]. In contrast to the case under the parabolic Hörmander condition, $\mathcal{P}_t f$ need not be differentiable in the direction V_0 however it is differentiable in the direction $\partial_t + V_0$. For more details see [32, Appendix A].

[F.3] For $f \in C_V^\infty(\mathbb{R}^N)$ we have that $(x,t) \mapsto \mathcal{P}_t f$ is smooth in both x and t, i.e. it is differentiable arbitrarily many times in every direction, see [30]. For $f \in C_b(\mathbb{R}^N)$ we may take a sequence $f_n \in C_V^\infty(\mathbb{R}^N)$ such that $\mathcal{P}_t f_n \in C_V^\infty(\mathbb{R}^N)$ and for each compact set $K \subseteq \mathbb{R}^N$ we have that $\mathcal{P}_t f_n$ and $V_{[\alpha_1]} \ldots V_{[\alpha_k]} \mathcal{P}_t f_n$ converge uniformly over K as n tends to ∞ to $\mathcal{P}_t f$ and $V_{[\alpha]} \ldots V_{[\alpha_k]} \mathcal{P}_t f$ respectively for each $k \in \mathbb{N}, \alpha_1, \ldots, \alpha_k \in \mathcal{A}$. We shall denote the space of all such functions by $\mathcal{D}_V^{2,\infty}(\mathbb{R}^N)$. See [32, Appendix A] for more details.

A.1.5 Miscellaneous technical facts

Lemma A.1.13. *Let X and Y be as in Example 2.3.2 then the vector fields $\{X, Y\}$ do not satisfy the UFG condition, in the sense that whether we take $X = V_0$ and $Y = V_1$ or viceversa, the UFG condition is not satisfied.*

Proof of Lemma A.1.13. In the definition of UFG condition take $Y = V_0$ and $X = V_1$ (the other case is simple to show) and assume that the UFG condition holds for some $m \in \mathbb{N}$. Denote by ad_X the map which takes a vector field Z to $[X, Z]$, then note that

$$(\mathrm{ad}_X)^k Y = \psi^{(k)}(x)\partial_y. \tag{A.5}$$

Here $\psi^{(k)}$ denotes the k^{th} derivative of ψ. Now $(\mathrm{ad}_X)^k Y$ commutes with the vector field Y and hence the only non-trivial vector fields in $\mathcal{R}_m$ are X, Y and $(\mathrm{ad}_X)^k Y$ for any $k \in \mathbb{N}$. By the UFG condition there exist smooth functions $\varphi_X, \varphi_{Y,k}$ such that

$$(\mathrm{ad}_X)^{m+1} Y = \sum_{k=0}^{m} \varphi_{Y,k} (\mathrm{ad}_X)^k Y + \varphi_X X.$$

We may write this as follows using (A.5)

$$\psi^{(m+1)}\partial_y = \sum_{k=0}^{m} \varphi_{Y,k} \psi^{(k)} \partial_y + \varphi_X \partial_x.$$

By considering the direction ∂_x we have that $\varphi_X = 0$, therefore we have

$$\psi^{(m+1)} = \sum_{k=0}^{m} \varphi_{Y,k} \psi^{(k)}.$$

Also note that since $\psi(x) = 0$ for all $x < 0$ we have that $\psi^{(k)}(x) = 0$ for all $x < 0$ and $k \in \mathbb{N}$; as ψ is smooth this gives that $\psi^{(k)}(0) = 0$ for all $k \in \mathbb{N}$. In particular, ψ solves the following initial value problem

$$\psi^{(m+1)}(x) = \sum_{k=0}^{m} \varphi_{Y,k}(x,y)\psi^{(k)}(x), \qquad \text{for all } x \geq 0$$

$$\psi^{(k)}(0) = 0, \qquad \text{for all } k \in \{0, 1, 2, \ldots, m\}.$$

However since ψ is smooth and the functions $\{\varphi_{Y,k}\}_{k \geq 0}$ are smooth, there is a (at east locally) unique solution to this initial value problem; the function which is constantly zero clearly satisfies the initial value problem. Therefore we have that $\psi \equiv 0$ (in a neighbourhood of zero), which gives a contradiction and hence the UFG condition is not satisfied. □

Lemma A.1.14. *Assume that the vector fields $V_0, \ldots, V_d$ satisfy the UFG condition. Let $\mathscr{S}$ be a maximal integral submanifold of $\hat{\Delta}_0$ and let $x, y \in \mathscr{S}$. Assume that x, y lie in the same coordinate neighbourhood $\mathscr{U}_{x_0}$ of a coordinate transformation Φ_{x_0} constructed in Section 3.1.2 . Then x and y lie in the same maximal integral submanifold of $\hat{\Delta}$ if and only if $\Phi_{x_0}^{n+1}(x) = \Phi_{x_0}^{n+1}(y)$.*

Proof of Lemma A.1.14. Assume that x, y both lie in the same maximal integral submanifold S of $\hat{\Delta}$. Then there is a time $T > 0$ and a path $p : [0, T] \to S$ satisfying the following ODE

$$\dot{p}(t) = \sum_{\alpha \in \mathcal{A}_m} V_{[\alpha]}(p(t))\psi_\alpha(t), \quad p(0) = x, \quad p(T) = y$$

for some piecewise linear input functions $\psi_\alpha : [0, T] \to \mathbb{R}$.

Now let $\tilde{p}(t) = \Phi_{x_0}(p(t))$ and let $\tilde{V}$ denote the representation of V in the coordinates defined by Φ_{x_0}, then we have

$$\dot{\tilde{p}}(t) = \sum_{\alpha \in \mathcal{A}_m} \tilde{V}_{[\alpha]}(\tilde{p}(t))\psi_\alpha(t).$$

Now by the properties in Proposition 3.1.9 we have that $\tilde{V}_{[\alpha]}^{n+1} = 0$ for all $\alpha \in \mathcal{A}$, and hence

$$\Phi_{x_0}^{n+1}(y) = \tilde{p}^{n+1}(T) = \tilde{p}^{n+1}(0) = \Phi_{x_0}^{n+1}(x).$$

Now assume that $\Phi_{x_0}^{n+1}(x) = \Phi_{x_0}^{n+1}(y)$.

Let $\tilde{\gamma}$ be any smooth curve that is contained in $(\Phi_{x_0}(\mathscr{U}_{x_0})) \cap (\mathbb{R}^n \times \{\Phi_{x_0}^{n+1}(x)\})$, and let $\dot{\tilde{\gamma}}(0) = \tilde{v}$. Define $\gamma = \Phi_{x_0}^{-1}(\tilde{\gamma})$ and $v = \dot{\gamma}(0)$. Now we have

$$\dot{\gamma}(0) = \mathscr{J}_z \Phi_{x_0}^{-1}(\gamma(0))\dot{\tilde{\gamma}}(0) = \mathscr{J}_z \Phi_{x_0}^{-1}(\gamma(0))\tilde{v}.$$

Since $\tilde{\gamma}$ is contained within $\mathbb{R}^n \times \{\Phi_{x_0}^{n+1}(x)\}$ we have that $\tilde{v} \in \mathbb{R}^n \times \{0\}$ and hence $v \in \hat{\Delta}(\gamma(0))$. Therefore the tangent space to $\Phi_{x_0}^{-1}(\mathrm{Im}(\Phi_{x_0}) \cap (\mathbb{R}^n \times \{\Phi_{x_0}^{n+1}(x)\}))$ at each point x' in this set is $\hat{\Delta}(x')$. Therefore $\Phi_{x_0}^{-1}(\mathrm{Im}(\Phi_{x_0}) \cap (\mathbb{R}^n \times \{\Phi_{x_0}^{n+1}(x)\})) \subseteq S_x$, where S_x is the maximal integral submanifold of $\hat{\Delta}$ which passes through x. In particular, we have that $y \in S_x$ as required. □

Lemma A.1.15. *Assume the vector fields $V_0, \ldots, V_d$ satisfy the UFG condition. Let $x, y \in \mathbb{R}^N$ be connected by an integral curve of one of the vector fields $V_{[\alpha]}, \alpha \in \mathcal{A}_m$; that is, $y = e^{TV_{[\alpha]}}(x)$ for some $T > 0$ and $\alpha \in \mathcal{A}_m$. Then, for all $h \in \mathcal{D}_V^{2,\infty}(\mathbb{R}^N)$,[29] we have*

$$h(y) - h(x) = \int_0^T (V_{[\alpha]}h)(\gamma(s))ds.$$

Proof of Lemma A.1.15. Let $\gamma(t) = e^{tV_{[\alpha]}}(x)$. Note that γ is continuous, hence the image of the interval $[0,T]$ through γ, $\gamma([0,T])$, is a compact set; in particular, there is some $r > 0$ such that $\gamma([0,T]) \subseteq B_r$. Then by Appendix A.1.4 **[F.3]** we have a sequence $\{h_n\}_{n\in\mathbb{N}} \subseteq C_V^{\infty}(\mathbb{R}^N)$ such that h_n converges uniformly to h on B_r and $V_{[\alpha]}h_n$ converges uniformly to $V_{[\alpha]}h$ on B_r, and therefore pointwise. Now using the chain rule we have

$$\begin{aligned}\frac{d}{ds}h_n(\gamma(s)) &= \dot{\gamma}(s) \cdot \nabla h_n(\gamma(s)) \\ &= V_{[\alpha]}(\gamma(s)) \cdot \nabla h_n(\gamma(s)) \\ &\overset{(1.3)}{=} V_{[\alpha]}h_n(\gamma(s)).\end{aligned}$$

[29]We recall that the set $\mathcal{D}_V^{2,\infty}(\mathbb{R}^N)$ has been introduced in Appendix A.1.4 **[F.3]**.

Integrating from 0 to T, and using that $\gamma(0) = x$ and $\gamma(T) = y$ we have

$$h_n(y) - h_n(x) = \int_0^T (V_{[\alpha]} h_n)(\gamma(s))ds.$$

Letting n tend to ∞ the statement follows. $\square$

Lemma A.1.16. *With notation of Section 5.1, suppose Hypothesis 5.1.1* **[H.1]** *holds. For any $g \in C_b(\mathbb{R}^n)$ define the functions $f(z, \zeta) = g(z)$ and $v_s(z, t) := \mathcal{P}_t f(z, \zeta_{s-t})$. Then v_s is smooth as a map from $\mathbb{R}^n \times (0, \infty)$ to $\mathbb{R}$, moreover it satisfies*

$$\partial_t v_s(z, t) = U_0 v_s(z, t) + \sum_{i=1}^{d} U_i^2 v_s(z, t). \tag{A.6}$$

Proof of Lemma A.1.16. Note that $z \in \mathbb{R}^n \mapsto v_s(z, t)$ is smooth (in any direction in $\mathbb{R}^n$) for each fixed $t > 0$ since $\mathcal{P}_t f$ is differentiable in all the directions spanned by $V_{[\alpha]}$ for all α (which span $\mathbb{R}^n$). To see that $t \mapsto v_s(z, t)$ is differentiable we first consider the case when f belongs to $C_V^\infty(\mathbb{R}^{n+1})$ then $t \mapsto \mathcal{P}_t f$ is differentiable (see Appendix A.1.4 **[F.3]**) and hence $t \mapsto v_s(z, t)$ is differentiable. Moreover using (A.35) we may differentiate v_s to find

$$\begin{aligned}
\partial_t v_s(z, t) &= V_0 \mathcal{P}_t f(z, \zeta_{s-t}) + \sum_{i=1}^{d} V_i^2 \mathcal{P}_t f(z, \zeta_{s-t}) - W_0(\zeta_{s-t}) \partial_\zeta \mathcal{P}_t f(z, \zeta_{s-t}) \\
&= V_0 \mathcal{P}_t f(z, \zeta_{s-t}) + \sum_{i=1}^{d} V_i^2 \mathcal{P}_t f(z, \zeta_{s-t}) - V_0^{(\perp)} \mathcal{P}_t f(z, \zeta_{s-t}) \\
&= (V_0 - V_0^{(\perp)}) \mathcal{P}_t f(z, \zeta_{s-t}) + \sum_{i=1}^{d} V_i^2 \mathcal{P}_t f(z, \zeta_{s-t}).
\end{aligned}$$

Now using the equality $V_0^{(\hat{\Delta})} = V_0 - V_0^{(\perp)}$ (see (1.12)), we have

$$\partial_t v_s(z, t) = V_0^{(\hat{\Delta})} v_s(z, t) + \sum_{i=1}^{d} V_i^2 v_s(z, t).$$

Note that, as differential operators, $V_0^{(\hat{\Delta})} = U_0$ and $V_i = U_i$ therefore we have that v_s satisfies (A.6).

To extend the proof to the case when f belongs to $C_b(\mathbb{R}^{n+1})$ we apply the argument of [32, Appendix A], so we only sketch this part of the proof. By Appendix

A.1.4 **[F.3]** if $f \in C_b(\mathbb{R}^{n+1})$ may take a sequence $f_n \in C_V^\infty(\mathbb{R}^{n+1})$ such that f_n converges to f and $V_{[\alpha_1]} \dots V_{[\alpha_k]} \mathcal{P}_t f_n$ converges uniformly on compacts of $\mathbb{R}^{n+1} \times (0, \infty)$ to $V_{[\alpha_1]} \dots V_{[\alpha_k]} \mathcal{P}_t f$ for any $k \geq 1$, and $\alpha_1, \dots, \alpha_k \in \mathcal{A}_m$. By the above argument we have

$$\partial_t(\mathcal{P}_t f_n(z, \zeta_{s-t})) = U_0 \mathcal{P}_t f_n(z, \zeta_{s-t}) + \sum_{i=1}^d U_i^2 \mathcal{P}_t f_n(z, \zeta_{s-t}).$$

For any $h > 0$ we have

$$\frac{\mathcal{P}_{t+h} f_n(z, \zeta_{s-(t+h)}) - \mathcal{P}_t f_n(z, \zeta_{s-t})}{h} = \frac{1}{h} \int_t^{t+h} U_0 \mathcal{P}_r f_n(z, \zeta_{s-r}) + \sum_{i=1}^d U_i^2 \mathcal{P}_r f_n(z, \zeta_{s-r}) dr\,;$$

therefore, letting n tend to ∞, we obtain

$$\frac{\mathcal{P}_{t+h} f(z, \zeta_{s-(t+h)}) - \mathcal{P}_t f(z, \zeta_{s-t})}{h} = \frac{1}{h} \int_t^{t+h} U_0 \mathcal{P}_r f(z, \zeta_{s-r}) + \sum_{i=1}^d U_i^2 \mathcal{P}_r f(z, \zeta_{s-r}) dr.$$

Letting now h tend to 0 we have that $(\mathcal{P}_t f)(z, \zeta_{s-t})$ is differentiable with respect to t and moreover

$$\partial_t(\mathcal{P}_t f(z, \zeta_{s-t})) = U_0 \mathcal{P}_t f(z, \zeta_{s-t}) + \sum_{i=1}^d U_i^2 \mathcal{P}_t f(z, \zeta_{s-t}).$$

That is, v_s is differentiable in both z and t as a map from $\mathbb{R}^n \times (0, \infty)$ to $\mathbb{R}$ and satisfies (A.6).

$\square$

Lemma A.1.17. *We use the notation of Section 5.2. If the map W^∞ is well defined on S_{x_0} (in the sense that $S_{x_0} \subseteq \mathrm{Dom}(W^\infty)$) and it is continuous when restricted to S_{x_0}, then W^∞ is also well defined and continuous on $\mathscr{S}_{x_0}$.*

Proof of Lemma A.1.17. First note that given any point $x \in \mathscr{S}_{x_0}$ we can find some $s \in \mathbb{R}$ and $z \in S_{x_0}$ such that $x = e^{sV_0^{(\perp)}}(z)$, in which case we have

$$W^\infty(x) = \lim_{t \to \infty} e^{tV_0^{(\perp)}}(x) = \lim_{t \to \infty} e^{(t+s)V_0^{(\perp)}}(z) = W^\infty(z). \tag{A.7}$$

Now $W^\infty(z)$ is well defined by assumption and hence $W^\infty(x)$ is well-defined.

To show that W^∞ is continuous on $\mathscr{S}_{x_0}$ take $\{x_k\}_k \subseteq \mathscr{S}_{x_0}$ and $x \in \mathscr{S}_{x_0}$ such that $x_k \to x$ as k tends to ∞, then we must show that $W^\infty(x_k)$ converges to $W^\infty(x)$

as k tends to ∞. Let $x_k = e^{s_k V_0^{(\perp)}}(z_k)$ and $x = e^{sV_0^{(\perp)}}(z)$ for some $s_k, s \in \mathbb{R}$ and $z_k, z \in S_{x_0}$. Without loss of generality we may assume that $s = 0$, otherwise consider the sequence $y_k := e^{-sV_0^{(\perp)}}(x_k)$.

Recall from Section 3.1.2 that we may take a local neighbourhood U_x of x and a coordinate transformation Φ. Then for k sufficiently large we have that $x_k \in U_x$, and hence $\Phi(x_k)$ converges to $\Phi(x)$. By the uniqueness of integral curves we have that

$$\Phi(e^{sV_0^{(\perp)}}(y)) = e^{s\tilde{V}_0^{(\perp)}}(\Phi(y)).$$

Recall that $\widetilde{V_0^{(\perp)}}$ denotes the representation of $V_0^{(\perp)}$ in the coordinates defined by Φ. Therefore

$$\Phi(x_k) = \Phi(e^{s_k V_0^{(\perp)}}(z_k)) = e^{s_k \tilde{V}_0^{(\perp)}}(\Phi(z_k)).$$

Since $\widetilde{V_0^{(\perp)}}$ only acts on the last coordinate we have that the first n components of $\Phi(e^{s_k V_0^{(\perp)}}(z_k))$ are equal to the first n components of $\Phi(z_k)$. In particular, the first n components of $\Phi(z_k)$ converge to the first n components of $\Phi(z)$. Now z_k and z lie on the same integral submanifold of $\hat{\Delta}$ and hence by Lemma A.1.14 the last component of $\Phi(z_k)$ is equal to the last coordinate of $\Phi(z)$. Therefore $\Phi(z_k)$ converges to $\Phi(z)$, and since Φ is a diffeomorphism we have that z_k converges to z.

Now since W^∞ is continuous on S_{x_0} and using (A.7) we have

$$W^\infty(x_k) = W^\infty(z_k) \to W^\infty(z) = W^\infty(x).$$

Therefore W^∞ is continuous on $\mathscr{S}_{x_0}$. $\square$

A.2 Proofs

This appendix contains all the proofs that we omitted in the main text.

Proof of Lemma 2.2.6. In [32], the authors proved estimates of the type (2.11) for first order derivatives of the semigroup $\mathcal{P}_t$. More precisely, they show that if (2.8) is satisfied, then (2.9) holds. This statement is the analogous of such results for second order derivatives and can be proved with the same procedures presented in [32]. Notice indeed that in [32, comments after Corollary 4.9] the authors explicitly observe how the technique used in that paper can be extended to cover derivatives of any order. $\square$

A.2.1 Proofs of Section 3.1 and Section 3.2

Proof of Lemma 3.1.1. First we show that $\mathrm{span}(\mathcal{R}_m)$ is contained in $\hat{\Delta}$. By definition $\hat{\Delta}$ contains $V_1, \dots, V_d$ and is invariant under $V_0, \dots, V_d$, hence by Note 2.3.5 we have that $V_{[\alpha]} \in \hat{\Delta}$ for all $\alpha \in \mathcal{A}_m$. By linearity we have that $\mathrm{span}(\mathcal{R}_m) \subseteq \hat{\Delta}$. We show that $\hat{\Delta}$ is contained in $\mathrm{span}(\mathcal{R}_m)$. It is sufficient to show that $\mathrm{span}(\mathcal{R}_m)$ contains $V_1, \dots, V_d$ and is invariant under $V_0, V_1, \dots, V_d$. Since $V_1, \dots, V_d \in \mathcal{R}_m$ it suffices to show that every vector field in $\mathrm{span}(\mathcal{R}_m)$ is invariant under $V_0, V_1, \dots, V_d$. Every vector field V in $\mathrm{span}(\mathcal{R}_m)$ can be locally expressed in the form

$$V = \sum_{\alpha \in \mathcal{A}_m} \varphi_\alpha V_{[\alpha]} \tag{A.8}$$

for some smooth functions φ_α. Therefore, again by Note 2.3.5, it is sufficient to show that $[V, V_j] \in \Delta_{\mathcal{R}_m}$ for V given by (A.8) and $j \in \{0, 1, \dots, d\}$. Note that

$$[V, V_j] = \sum_{\alpha \in \mathcal{A}_m} [\varphi_\alpha V_{[\alpha]}, V_j] = \sum_{\alpha \in \mathcal{A}_m} \varphi_\alpha [V_{[\alpha]}, V_j] - V_j(\varphi_\alpha) V_{[\alpha]}.$$

Now $[V_{[\alpha]}, V_j]$ and $V_{[\alpha]}$ are in $\mathrm{span}(\mathcal{R}_m)$ and hence $\mathrm{span}(\mathcal{R}_m)$ is invariant under $V_0, V_1, \dots, V_d$. Therefore $\hat{\Delta} = \mathrm{span}(\mathcal{R}_m)$; similarly one can show that $\hat{\Delta}_0 = \mathrm{span}(\mathcal{R}_{m,0})$. $\square$

Proof of Proposition 3.1.9. Let us start by proving i). Construct Φ as described before the statement of Proposition 3.1.9. After the change of coordinates Φ the vector V is expressed as

$$\tilde{V}(\mathbf{z}) = [(\mathcal{J}_x \Phi) \cdot V(x)]\,|_{x=\Phi^{-1}(\mathbf{z})}\,. \tag{A.9}$$

As we have already observed the last $N - n$ rows of the Jacobian matrix $\mathcal{J}\Phi$ are orthogonal to vectors in Δ, see (3.6). Since $V \in \Delta$, the statement follows.

To prove ii), we first observe that by i), the vector fields $\{\partial_{z_j}\}_{j=1}^n$ belong to Δ. Moreover, by Note 2.3.5, we have that $[\tilde{W}, \partial_{z_j}] \in \Delta$, for all $j = 1, \dots, n$. The field

$[\tilde{W}, \partial_{z_j}]$ can be calculated explicitly:

$$[\tilde{W}, \partial_{z_j}] = \left[\sum_{i=1}^{N} \tilde{W}^i \partial_{z_i}, \partial_{z_j}\right] = -\sum_{i=1}^{N} \frac{\partial \tilde{W}^i}{\partial z_j} \partial_{z_i} .$$

Because $[\tilde{W}, \partial_{z_j}] \in \Delta$, one must have

$$\frac{\partial \tilde{W}^i}{\partial z_j} = 0 \quad \text{for all } j = 1, \ldots, n, i = n+1, \ldots, N .$$

This concludes the proof. $\square$

Proof of Lemma 3.1.16. Since $S \subseteq \mathscr{S}$ we have that $\overline{S} \subseteq \overline{\mathscr{S}}$, therefore it is sufficient to show that if $x \in \partial S$ then $x \notin \mathscr{S}$. Assume for a contradiction there exists some $x \in \partial S \cap \mathscr{S}$. Since $x \in \mathscr{S}$ there exists a neighbourhood $U \subseteq \mathscr{S}$ which contains x and on which the coordinate transformation Φ constructed at the beginning of Section 3.1.2 is well defined. Now $x \in \partial S$ implies there exists a sequence $\{x_k\} \subseteq S$ such that x_k converges to x. For k sufficiently large x_k belongs to U and, since the coordinate transformation is smooth, we have

$$\Phi^{n+1}(x) = \lim_{k \to \infty} \Phi^{n+1}(x_k) , \tag{A.10}$$

having used the notation 3.1.2. However x_k all belong to the same maximal integral submanifold of $\hat{\Delta}$ and hence $\Phi^{n+1}(x_k)$ is constant for k large enough, by Lemma A.1.14. However this implies, for k large enough, that $\Phi^{n+1}(x) = \Phi^{n+1}(x_k)$; so by (A.10) and Lemma A.1.14 we have that x and x_k lie in the same maximal integral submanifold of $\hat{\Delta}$. However this gives a contradiction, since $x_k \in S$ and $x \notin S$. $\square$

Proof of Lemma 3.1.17. We will prove that the set of points $K := \{x \in \mathscr{S}_{x_0} : V_0^{(\perp)}(x_0) = 0\} \subseteq \mathscr{S}_{x_0}$ is both open and closed in (the topology of) $\mathscr{S}_{x_0}$, hence it has to be the whole manifold $\mathscr{S}_{x_0}$ – see **[SA.2]** and Appendix A.1.3 for clarifications on the manifold topology. Such a set is clearly closed (in $\mathbb{R}^N$ and hence in the manifold topology) as it is the intersection between $\mathscr{S}_{x_0}$ and the preimage of 0 through a continuous function. To prove that it is also open, we will show that for any $x \in K$ there exists an open neighbourhood of x, O_x, which is contained in K. Let $x \in \mathbb{R}^N$ such that $V_0^{(\perp)}(x) = 0$ and let $n = n(x)$ be the rank of $\hat{\Delta}_0$ at x; then there exist n vectors in $\hat{\Delta}_0(x)$ which span $\hat{\Delta}_0$ at x. Notice that, by construction, such vectors must belong to $\hat{\Delta}(x)$, as by Lemma 3.1.1 $\hat{\Delta}(x) = \hat{\Delta}_0(x)$ if $V_0^{(\perp)}(x) = 0$. By the smoothness of the vector fields and because x is a regular point for both distributions, there exists a neighbourhood O_x of x such that the same n vectors span $\hat{\Delta}_0(y)$ for every $y \in O_x$. Because $V_0^{(\perp)}(y)$ is orthogonal to all the vectors in $\hat{\Delta}_0(y)(= \hat{\Delta}(y))$, it must be the case that $V_0^{(\perp)}(y) = 0$ on O_x (otherwise the rank of $\hat{\Delta}_0$ would increase, which is impossible as the rank stays constant on the orbits). Therefore $O_x \subseteq K$ and the proof is concluded. $\square$

Proof of Proposition 3.2.1. We emphasize that this proof heavily relies on the fact that the integral manifolds of $\hat{\Delta}_0$ coincide with the orbits of $\hat{\Delta}_0$, see Proposition 3.1.3.

• *Proof of i).* Let $\mathscr{S}$ be one of the integral manifolds of $\hat{\Delta}_0$ and suppose $x \in \partial\mathscr{S}$. To begin with, we show that $\mathscr{S}_x \subseteq \bar{\mathscr{S}}$. To this end, let y be any point in $\mathscr{S}_x$. We want to show that $y \in \bar{\mathscr{S}}$. By Proposition 3.1.3 the integral manifold $\mathscr{S}_x$ is given by the orbit through x of the vector fields in $\hat{\Delta}_0$, and hence y can be written as the end point of a curve which starts from x and is a piecewise integral curve for vector fields in $\hat{\Delta}_0$. By considering each piece of the integral curve separately, if needed, we may assume that $y = e^{TV}(x)$ for some $T > 0$ and $V \in \hat{\Delta}_0$. Since $x \in \overline{\mathscr{S}}$, there is a sequence $\{x_k\}_k$ converging to x and such that $\{x_k\}_k \subseteq \mathscr{S}$. Set $y_k := e^{TV}(x_k)$ and note that $\{y_k\}_k$ belongs to $\mathscr{S}$ since $\mathscr{S}$ is an orbit of $\hat{\Delta}_0$. We have that y_k converges to y since the map $z \mapsto e^{TV}(z)$ is continuous. Therefore $y \in \overline{\mathscr{S}}$ which implies that $\mathscr{S}_x \subseteq \overline{\mathscr{S}}$ as y is an arbitrary point in $\mathscr{S}_x$. However $\mathscr{S}$ and $\mathscr{S}_x$ are both maximal integral submanifolds so they are either disjoint or they coincide; since $x \in \mathscr{S}_x$ and $x \notin \mathscr{S}$ they must be disjoint, hence $\mathscr{S}_x \subseteq \partial\mathscr{S}$.

• *Proof of ii).* Note that by the Stroock and Varadhan Theorem, Theorem 3.1.4, we have $\mathbb{P}_x(X_t \in \overline{\mathscr{S}_x}) = 1$ for any $x \in \mathbb{R}^N$. From the reasoning in the proof of point i), we know that if $x \in \partial\mathscr{S}$ then $\mathscr{S}_x \subseteq \partial\mathscr{S}$, so that $\overline{\mathscr{S}_x} \subseteq \partial\mathscr{S}$. Therefore for any $x \in \partial\mathscr{S}$ we have

$$\mathbb{P}_x(X_t \in \partial\mathscr{S}) \geq \mathbb{P}_x(X_t \in \overline{\mathscr{S}_x}) = 1.$$

$\square$

Proof of Proposition 3.2.3. Here we consider the case $V_0^{(\perp)}(x_0) \neq 0$ (which, by

Lemma 3.1.17, implies $V_0^{(\perp)}(x) \neq 0$ for every $x \in \mathscr{S}_{x_0}$). Consider the control problem (3.2) associated with the SDE (1.1). If we can show that any solution $p(t)$ of (3.2) has the property that $p(t) \in S_{e^{tV_0^{(\perp)}}(x_0)}$ for all t, the result then follows by the Stroock and Varadhan Support Theorem.[30] Let us now define the set

$$C := \left\{ t \in \mathbb{R} : p(t) \in S_{e^{tV_0^{(\perp)}}(x_0)} \right\}.$$

Note that C is non-empty since $0 \in C$; if we can show that C is open and closed as a subset of $\mathbb{R}$ then we must have that $C = \mathbb{R}$ which implies the desired result. Let us start by showing that C is open in $\mathbb{R}$. To this end, fix an arbitrary point $t_0 \in C$; without loss of generality we may assume that $t_0 = 0$ (otherwise we consider the path $q(t) := p(t + t_0)$). We will show that there exists an open neigbourhood of 0 which is contained in C. To show this fact we will make use of the (local) change of coordinates defined in Section 3.1.2. Let $\Phi_{x_0} : \mathscr{U}_{x_0} \to \tilde{\mathscr{U}}_{x_0}$ be the coordinate transformation defined on a local neighbourhood of x_0. Take $\varepsilon > 0$ sufficiently small that $p(t) \in \mathscr{U}_{x_0}$ for all $t \leq \varepsilon$. It is sufficient to show that $p(t) \in S_{e^{tV_0^{(\perp)}}(x_0)}$ for all $t \in (-\varepsilon, \varepsilon)$. Let $\tilde{p}(t) = \Phi_{x_0}(p(t))$; consistently with the notation set in Section 2.1, we shall denote the first n components of $\tilde{p}(t)$ by $z(t)$, the $(n+1)^{th}$ component by $\zeta(t)$ and the last $N - (n+1)$ components by $a(t)$. That is, $\tilde{p}(t) = (z(t), \zeta(t), a(t)) = \Phi_{x_0}(p(t))$; hence, in particular,

$$\zeta(t) = \Phi_{x_0}^{n+1}(p(t)). \tag{A.11}$$

By Lemma A.1.14 and (A.11), to prove that $p(t) \in S_{e^{tV_0^{(\perp)}}(x_0)}$ it is sufficient to show that the following holds

$$\zeta(t) = \Phi_{x_0}^{n+1}(e^{tV_0^{(\perp)}}(x_0)). \tag{A.12}$$

Differentiating the equation $\tilde{p}(t) = \Phi_{x_0}(p(t))$ with respect to t and using (3.2) and (A.9), we see that $\tilde{p}$ satisfies the equation

$$\frac{d\tilde{p}(t)}{dt} = \tilde{V}_0(\tilde{p}(t)) + \sqrt{2} \sum_{i=1}^{d} \tilde{V}_i(\tilde{p}(t))\psi_i.$$

[30] Note that by Theorem 3.1.4 the path $\{X_t^{(x_0)}(\omega)\}_{t\in[0,T]}$ is a limit, in $C([0,T], \|\cdot\|_\infty)$, of solutions to the control problem (3.2). Because uniform convergence implies pointwise convergence, for each fixed $t \geq 0$ the point X_t is a limit of $\{p(t) : p \text{ is a solution to (3.2)}\}$.

Since $V_i \in \hat{\Delta}$ for $i = 1, \dots, d$ and $\hat{\Delta}$ is invariant under V_0 we have, using Proposition 3.1.9 (ii) and the notation (2.2)-(2.3),

$$\begin{aligned}
\frac{dz(t)}{dt} &= U_0(z(t), \zeta(t), a(t)) + \sqrt{2}\sum_{i=1}^{d} U_i(z(t), \zeta(t), a(t))\psi_i, \\
\frac{d\zeta(t)}{dt} &= W_0(\zeta(t), a(t)), \\
\frac{da(t)}{dt} &= 0.
\end{aligned} \tag{A.13}$$

(The above is completely analogous to what we have done to obtain (3.7)-(3.9)). From equation (A.13) we then have

$$\zeta(t) = e^{tW_0}(\zeta(0)) \overset{(A.11)}{=} e^{tW_0}(\Phi^{n+1}_{x_0}(x_0)).$$

In order to prove (A.12) it remains to show that $e^{tW_0}(\Phi^{n+1}_{x_0}(x_0)) = \Phi^{n+1}_{x_0}(e^{tV_0^{(\perp)}}(x_0))$. By uniqueness of the integral curves, to prove this equality we must show

$$\frac{d}{dt}\Phi^{n+1}_{x_0}(e^{tV_0^{(\perp)}}(x_0)) = W_0(\Phi^{n+1}_{x_0}(e^{tV_0^{(\perp)}}(x_0))).$$

This follows from

$$\begin{aligned}
\frac{d}{dt}(\Phi^{n+1}_{x_0}(e^{tV_0^{(\perp)}}(x_0))) &= \nabla_x \Phi^{n+1}_{x_0} x(e^{tV_0^{(\perp)}}(x_0))V_0^{(\perp)}(e^{tV_0^{(\perp)}}(x_0)) \\
&= W_0(\Phi^{n+1}_{x_0}(e^{tV_0^{(\perp)}}(x_0))),
\end{aligned}$$

where the last equality is a consequence of (A.9) and of the fact that $\tilde{V_0}^{(\perp)} = (0, \dots, 0, W_0, 0, \dots, 0)$ (see comments after (3.7)-(3.9)). This proves (A.12), so C is open. Now we show that C is closed by showing that $\mathbb{R} \setminus C$ is open.

Assume there exists t_0 such that $p(t_0) \notin S_{e^{t_0 V_0^{(\perp)}}(x_0)}$. Now we may take ε sufficiently small that $p(t) \in \mathscr{U}_{p(t_0)}$ whenever $|t - t_0| < \varepsilon$. It is sufficient to show that $p(t) \notin C$ whenever $|t - t_0| < \varepsilon$. For a contradiction assume that there exists some $\bar{t} \in (t_0 - \varepsilon, t_0 + \varepsilon)$ with $\bar{t} \in C$ i.e. $p(\bar{t}) \in S_{e^{\bar{t}V_0^{(\perp)}}(x_0)}$; then by the same argument as above, we have that $p(t_0) \in S_{e^{(t_0-\bar{t})V_0^{(\perp)}}(p(\overline{t_0}))}$. Now by Lemma 2.3.8 applied to the points $p(\bar{t})$ and $e^{\bar{t}V_0^{(\perp)}}(x_0)$ and to the vector field $V_0^{(\perp)}$ we have that

$$S_{e^{(t_0-\bar{t})V_0^{(\perp)}}(p(\bar{t}))} = S_{e^{(t_0-\bar{t})V_0^{(\perp)}}(e^{\bar{t}V_0^{(\perp)}}(x_0))} = S_{e^{t_0V_0^{(\perp)}}(x_0)}.$$

Therefore $p(t_0) \in S_{e^{t_0V_0^{(\perp)}}}$ and we have a contradiction since $t_0 \notin C$ therefore there is no $\bar{t} \in (t_0 - \varepsilon, t_0 + \varepsilon) \cap C$, hence $(t_0 - \varepsilon, t_0 + \varepsilon) \subseteq \mathbb{R} \setminus C$. That is, C is closed in $\mathbb{R}$ and we have $C = \mathbb{R}$ as required. $\square$

Proof of Proposition 3.2.7. For every $x \in \mathbb{R}^N$ and $t > 0$, let $g_t(x) := \mathbb{P}_x(X_t^{(x)} \notin \mathscr{S}_x)$ and notice that $E_t = \{x \in \mathbb{R}^N : \mathbb{P}_x(X_t^x \notin \mathscr{S}_x) > 0\} = \{x \in \mathbb{R}^N : g_t(x) > 0\}$. Suppose the SDE (1.1) admit an invariant measure, μ. Because $E = \cup_{t>0} E_t$, if we prove that $\mu(E_t) = 0$ for every $t \geq 0$, then it follows that $\mu(E) = 0$ (as $\{E_t\}_{t\geq 0}$ is an increasing sequence of sets). So we concentrate on proving the first statement. To this end, define $\mathscr{S}^\ell$ to be the union of all the maximal integral submanifolds of $\hat{\Delta}_0$ of dimension ℓ and notice that $\cup_{\ell=0}^N \mathscr{S}^\ell = \mathbb{R}^N$; moreover, for every (arbitrary but fixed) $t > 0$, set $E_t^\ell := \{x \in \mathscr{S}^\ell : \mathbb{P}_x(X_t^x \notin \mathscr{S}^\ell) > 0\}$. We now proceed in two steps.

- *Step 1*: show that
$$E_t = \bigcup_{\ell=0}^N E_t^\ell.$$

Note that $E_t^\ell \subseteq E_t$; indeed, if $x \in E_t^\ell$ then $\mathscr{S}_x \subseteq \mathscr{S}^\ell$ and $\mathbb{P}_x(X_t^x \notin \mathscr{S}_x) \geq \mathbb{P}_x(X_t^x \notin \mathscr{S}^\ell) > 0$. Therefore $x \in E_t$. It remains to show that if $x \in E_t$ then there exists some ℓ such that $x \in E_t^\ell$. Fix $x \in E_t$ and let ℓ denote the dimension of $\mathscr{S}_x$. By the Stroock and Varadhan Support Theorem (Theorem 3.1.4) we have that $\mathbb{P}_x(X_t^x \in \overline{\mathscr{S}_x}) = 1$ for every $t \geq 0$, hence

$$\mathbb{P}_x(X_t^x \in \partial\mathscr{S}_x) = \mathbb{P}_x(X_t^x \notin \mathscr{S}_x) = g_t(x).$$

By Proposition 3.2.1 we have that $\partial\mathscr{S}_x$ is contained in the set $\cup_{k<\ell}\mathscr{S}^k$. In particular, we have that $\partial\mathscr{S}_x$ is disjoint from $\mathscr{S}^\ell$ and hence

$$g_t(x) = \mathbb{P}_x(X_t^x \in \partial\mathscr{S}_x) \leq \mathbb{P}_x(X_t^x \notin \mathscr{S}^\ell).$$

Since $x \in E_t$ we have that $g_t(x) > 0$ and therefore $\mathbb{P}_x(X_t^x \notin \mathscr{S}^\ell) > 0$, which, by definition, gives that $x \in E_t^\ell$.

• *Step 2*: show that $\mu(E_t^\ell) = 0$ for all $\ell \in \{0, \dots, N\}$. To this end, set $g_t^\ell(x) := \mathbb{P}_x(X_t^x \notin \mathscr{S}^\ell)$; then the set of $x \in \mathscr{S}^\ell$ such that $g_t^\ell(x) > 0$ is the set E_t^ℓ. Therefore it is sufficient to show that $\int_{\mathscr{S}^\ell} g_t^\ell(x)\mu(dx) = 0$ for all $\ell \in \{0, \dots, N\}$. Assume this is not the case; that is, assume there exists some $\bar{\ell}$ such that

$$\int_{\mathscr{S}^{\bar{\ell}}} \mathbb{P}_x(X_t^x \notin \mathscr{S}^{\bar{\ell}})\mu(dx) > 0. \tag{A.14}$$

We will let $\bar{\ell}$ be the maximum index such that (A.14) holds. Since μ is an invariant measure we have that

$$\mu(\mathscr{S}^{\bar{\ell}}) = \int_{\mathbb{R}^N} \mathbb{P}_x(X_t^x \in \mathscr{S}^{\bar{\ell}})\mu(dx) = \sum_{k=0}^{N} \int_{\mathscr{S}^k} \mathbb{P}_x(X_t^x \in \mathscr{S}^{\bar{\ell}})\mu(dx). \tag{A.15}$$

Fix $k \in \{0, \dots, N\}$ and first consider the case when $k > \bar{\ell}$. Since $\bar{\ell}$ was chosen to be maximal such that (A.14) holds we must have

$$\int_{\mathscr{S}^k} \mathbb{P}_x(X_t^x \notin \mathscr{S}^k)\mu(dx) = 0 \quad \text{if } k > \bar{\ell}.$$

This is equivalent to saying that the μ- measure of the set $E_t^k = \{x \in \mathscr{S}^k : \mathbb{P}_x(X_t^x \notin \mathscr{S}^k) > 0\}$ is zero. Since $k \neq \bar{\ell}$, we have $\{x \in \mathscr{S}^k : \mathbb{P}_x(X_t^x \in \mathscr{S}^{\bar{\ell}}) > 0\} \subseteq E_t^k$, so the μ- measure of the set $\{x \in \mathscr{S}^k : \mathbb{P}_x(X_t^x \in \mathscr{S}^{\bar{\ell}}) > 0\}$ is zero as well. Therefore

$$\int_{\mathscr{S}^k} \mathbb{P}_x(X_t \in \mathscr{S}^{\bar{\ell}})\mu(dx) = 0 \quad \text{for all } k > \bar{\ell},$$

so that

$$\sum_{k=\bar{\ell}+1}^{N} \int_{\mathscr{S}^k} \mathbb{P}_x(X_t^x \in \mathscr{S}^\ell)\mu(dx) = 0\,. \tag{A.16}$$

Now consider the case $k < \bar{\ell}$. In this case we have

$$\sum_{k=0}^{\bar{\ell}-1} \int_{\mathscr{S}^k} \mathbb{P}_x(X_t^x \in \mathscr{S}^{\bar{\ell}})\mu(dx) = 0\,, \tag{A.17}$$

as by Proposition 3.2.1 the dimension of the manifold in which X_t evolves can only either decrease or stay the same along the paths of the SDE. Putting together (A.15), (A.16) and (A.17), one has

$$\mu(\mathscr{S}^{\bar{\ell}}) = \int_{\mathscr{S}^{\bar{\ell}}} \mathbb{P}_x(X_t^x \in \mathscr{S}^{\bar{\ell}})\mu(dx).$$

Writing $\mathbb{P}_x(X_t^x \in \mathscr{S}^{\bar{\ell}})$ as $1 - \mathbb{P}_x(X_t^x \notin \mathscr{S}^{\bar{\ell}})$ we obtain

$$\mu(\mathscr{S}^{\bar{\ell}}) = \int_{\mathscr{S}^{\bar{\ell}}} \mathbb{P}_x(X_t \in \mathscr{S}^{\bar{\ell}})\mu(dx) = \mu(\mathscr{S}^{\bar{\ell}}) - \int_{\mathscr{S}^{\bar{\ell}}} \mathbb{P}_x(X_t \notin \mathscr{S}^{\bar{\ell}})\mu(dx),$$

which gives

$$\int_{\mathscr{S}^{\bar{\ell}}} \mathbb{P}_x(X_t \notin \mathscr{S}^{\bar{\ell}})\mu(dx) = 0.$$

This contradicts (A.14) and hence we must have that the statement holds. □

Proof of Lemma 3.2.9. Fix $x, y \in S$ and $f \in C_b(\mathbb{R}^N)$ and assume first that x, y are such that there exists a path $\gamma : [0, T] \to \mathbb{R}^N$ with $\gamma(0) = x, \gamma(T) = y$ and $\dot{\gamma}(t) = V_{[\alpha]}(\gamma(t))$, for some $\alpha \in \mathcal{A}_m$. Clearly the final time T will depend on x and y, $T = T_{x,y}$. By Lemma A.1.15 and by Appendix A.1.4 **[F.3]** we have

$$\mathcal{P}_t f(y) - \mathcal{P}_t f(x) = \int_0^{T_{x,y}} (V_{[\alpha]}\mathcal{P}_t f)(\gamma(s))ds.$$

Take a compact set K such that $K \supseteq \gamma([0, T_{x,y}])$, then by (2.9) we have

$$|\mathcal{P}_t f(y) - \mathcal{P}_t f(x)| \leq \sup_{x \in K}(V_{[\alpha]}\mathcal{P}_t f)(x)T_{x,y} \leq c_{t_0,K}^{\frac{1}{2}} e^{-\lambda(t-t_0)/2} T_{x,y}\|f\|_\infty. \tag{A.18}$$

Letting t tend to ∞ we obtain the result. For any $x, y \in S$ we can take a piecewise integral curve connecting x and y, hence applying the above argument to each piece of the curve we obtain (3.19). □

Proof of Proposition 3.2.10. Assume there exists an invariant measure μ with $\mu(S) = 1$; we must show that μ is the unique invariant measure such that $\mu(S) = 1$. Integrating (3.19) with respect to μ we obtain

$$\lim_{t\to\infty} \left| \mathcal{P}_t f(x) - \int_S \mathcal{P}_t f(y)\mu(dy) \right| = 0, \text{ for every } x \in S. \tag{A.19}$$

(Here exchanging the integral and limit is justified by the dominated convergence theorem and using Appendix A.1.5 **[F.1]**). The invariance of μ and (A.19) imply

(3.20). Because (3.20) holds for every $x \in S$, by the uniqueness of the limit we have that μ must be the only invariant measure supported on S.

It remains to show that μ is ergodic. Suppose there exists $t > 0$ and a Borel set $E \subseteq \mathbb{R}^N$ such that $\mathcal{P}_t \mathbb{1}_E = \mathbb{1}_E$ μ-almost everywhere. Then by the semigroup property we have for every $n \in \mathbb{N}$ $\mathcal{P}_{nt} \mathbb{1}_E = \mathbb{1}_E$ μ-almost everywhere. Now squaring and integrating (3.20) with respect to μ we have that $\mathcal{P}_t f$ converges to $\int_S f d\mu$ in L^2_μ for each $f \in C_b(\mathbb{R}^N)$. Then since $C_b(\mathbb{R}^N)$ is dense in L^2_μ (see [68, Theorem 3.14]), for all $f \in L^2_\mu$ we have

$$\lim_{n \to \infty} \left\| \mathcal{P}_{nt} f - \int_S f d\mu \right\|_{L^2_\mu} = 0.$$

By taking a subsequence if necessary we get convergence μ-almost everywhere. In particular, taking $f = \mathbb{1}_E$ we get $\mu(E) = \mathbb{1}_E(x)$ μ-almost everywhere. Hence $\mu(E) = 0$ or 1, and μ is ergodic. □

A.2.2 Proofs and auxiliary results of Chapter 4

Note that (2.8) is equivalent to the following condition:

$$\xi^T V_{[\alpha * 0]}(x) V_{[\alpha]}(x) \xi \leq -\lambda_0 \left| \xi^T V_{[\alpha]}(x) \right|^2, \quad \text{for every } x, \xi \in \mathbb{R}^N \text{ and } \alpha \in \mathcal{A}_m. \tag{A.20}$$

Lemma A.2.1. *Consider the following SDE in $\mathbb{R}^N$*

$$dX_t = V_0(X_t)dt + \sqrt{2} \sum_{k=1}^{N} e_i dB_t^i, \quad X_0 = x,$$

where $\{e_i\}$ are the canonical basis vectors of $\mathbb{R}^N$. If the Obtuse Angle Condition (A.20) *holds for the above SDE then V_0 is unbounded and X_t^i is independent of X_t^j for each $t > 0$ and $i \neq j$.*

Proof of Lemma A.2.1. In this case the OAC (A.20) becomes

$$\sum_{j=1}^{N} \partial_i V_0^j(x) \xi^i \xi^j \leq -\lambda_0 |\xi^i|^2, \quad \forall x, \xi \in \mathbb{R}^N, i \in \{1, \ldots, N\}.$$

Fix some $i \in \{1, \ldots, N\}$ and take $\xi = e_i$; then we have

$$\partial_i V_0^i(x) \leq -\lambda_0, \quad \forall x \in \mathbb{R}^N.$$

Integrating the above gives $V_0^i(x) \leq V_0^i(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_N) - \lambda_0 x_i$ for $x_i > 0$ and $i \in \{1, \ldots, N\}$. Now letting x_i tend to ∞ we must have that $V_0(x)$ is unbounded below.

Moreover, if we take $\xi = e_i + Ke_k$ for some $i \neq k$ then by (A.20) we have

$$\partial_i V_0^i + K\partial_i V_0^k(x) \leq -\lambda_0, \quad \forall x \in \mathbb{R}^N, i \in \{1, \ldots, N\}, k \neq i, K \in \mathbb{R}.$$

By considering both the cases when K is large and negative, and when K is large and positive we must have that $\partial_i V_0^k = 0$ for $k \neq i$. Therefore X_t^i is independent of X_t^j for $i \neq j$. □

Lemma A.2.2. *Consider the SDE* (4.1) *when* $N = 1$, *i.e. consider the SDE*

$$dX_t = V_{0,It\hat{o}}(X_t)dt + \sqrt{2}\sum_{k=1}^{d} V_k(X_t)dB_t^k. \tag{A.21}$$

Then we may find a vector field U_1 *such that* X_t *is a weak solution to the SDE*

$$dX_t = V_{0,It\hat{o}}(X_t)dt + \sqrt{2}U_1(X_t)dW_t \tag{A.22}$$

for some one-dimensional Brownian motion $\{W_t\}_{t\geq 0}$. *Moreover, if the Local Obtuse Angle Condition* (1.17) *is satisfied by the vector fields in* (A.21), *then we have*

$$U_1(x)[U_1, V_0](x) \leq -\lambda(x)U_1(x)^2, \quad \textit{for every } x \in \mathbb{R},$$

where V_0 *is defined by* (4.2).

Proof. Define the process

$$W_t = \sum_{i=1}^{d} \int_0^t \frac{V_i(X_s^x)}{\sqrt{\sum_{j=1}^{d} |V_j(X_s^{(x)})|^2}} dB_s^i.$$

By the Levy Characterisation of Brownian motion (see [60, Theorem 3.3.16]), W_t is a one-dimensional Brownian motion. With this in mind, we have

$$
\begin{aligned}
dX_t^{(x)} &= V_{0,\text{Itô}}(X_t^{(x)})dt + \sqrt{2}\sum_{i=1}^{d} V_i(X_t^{(x)})dB_t^i \\
&= V_{0,\text{Itô}}(X_t^{(x)})dt + \sqrt{2}\left(\sum_{j=1}^{d}|V_j(X_t^{(x)})|^2\right)^{\frac{1}{2}} dW_t \\
&= V_{0,\text{Itô}}(X_t^{(x)})dt + \sqrt{2}U_1(X_t^{(x)})dW_t,
\end{aligned}
$$

where we set

$$
U_1(x) := \left(\sum_{j=1}^{d}|V_j(x)|^2\right)^{\frac{1}{2}}.
$$

Since

$$
U_1(x)U_1'(x) = U_1(x)\sum_{j=1}^{d} V_j(x)V_j'(x)\left(\sum_{i=1}^{d}|V_i(x)|^2\right)^{-\frac{1}{2}} = \sum_{j=1}^{d} V_j(x)V_j'(x)
$$

we have that $X_t^{(x)}$ satisfies the Stratonovich SDE

$$
dX_t^{(x)} = V_0(X_t^{(x)})dt + \sqrt{2}U_1(X_t^{(x)}) \circ dW_t.
$$

Note that

$$
\begin{aligned}
U_1(x)[U_1, V_0](x) &= U_1(x)^2 V_0'(x) - U_1(x)U_1'(x)V_0(x) \\
&= \sum_{j=1}^{d}|V_j(x)|^2 V_0'(x) - V_j(x)V_j'(x)V_0(x) \\
&= \sum_{j=1}^{d} V_j(x)[V_j, V_0](x).
\end{aligned}
$$

Therefore, if (1.17) is satisfied, we have

$$
U_1(x)[U_1, V_0](x) \leq -\lambda(x)\sum_{j=1}^{d}|V_j(x)|^2 = -\lambda(x)|U_1(x)|^2. \tag{A.23}
$$

$\square$

Note that the transformation in Lemma A.2.2 does not necessarily preserve the UFG condition however it does preserve the property of being locally finitely generated.

Let us now recall that a one dimensional SDE with multiplicative noise can be recast into a (one-dimensional) SDE with additive noise by using a Lamperti transformation, see [60, Section 5.2.C], assuming the coefficients of the initial SDE are bounded and satisfy an ellipticity condition.

Lemma A.2.3. *Consider a one-dimensional SDE with multiplicative noise of the form* (A.22) *and suppose the vector field U_1 appearing in* (A.22) *is such that* (A.22) *is uniformly elliptic. Then we can construct a smooth diffeomorphism h such that $Y_t := h(X_t)$ is the solution to*

$$dY_t = b_Y(Y_t)dt + \sqrt{2}dB_t \tag{A.24}$$

for some smooth function b_Y. Moreover, (A.22) *satisfies the Obtuse Angle condition* (1.16) *with constant λ_0 if and only if $b'_Y \leq -\lambda_0$.*

Proof of Lemma A.2.3. Consider the one dimensional SDE in Itô form (A.22). By the uniform ellipticity assumption there is some constant $\nu > 0$ such that $U_1(x) \geq \nu$ for all $x \in \mathbb{R}$. Fix some arbitrary $x_0 \in \mathbb{R}$ and define the function h as follows

$$h(x) = \int_{x_0}^{x} \frac{1}{U_1(y)} dy.$$

Let $Y_t = h(X_t)$, then Y_t is a strong solution of the SDE (A.24) where

$$b_Y(y) = \frac{V_{0,\text{Itô}}(h^{-1}(y))}{U_1(h^{-1}(y))} - U_1'(h^{-1}(y)) \overset{(4.2)}{=} \frac{V_0(h^{-1}(y))}{U_1(h^{-1}(y))}.$$

The derivative of b_Y is given by

$$\begin{aligned} b_Y'(y) &= \frac{V_0'(h^{-1}(y))U_1(h^{-1}(y)) - U_1'(h^{-1}(y))V_0(h^{-1}(y))}{U_1(h^{-1}(y))^2} \frac{d}{dy} h^{-1}(y) \\ &= \frac{[U_1, V_0](h^{-1}(y))U_1(h^{-1}(y))}{U_1(h^{-1}(y))^2}. \end{aligned}$$

From the above the statement follows. □

Note that we can formulate the Local Obtuse Angle Condition for a general SDE which satisfies the UFG condition as the requirement that there is some measurable function $\lambda : \mathbb{R}^N \to \mathbb{R}$ such that for all f sufficiently smooth

$$([V_{[\alpha]}, V_0]f)(x)(V_{[\alpha]}f)(x) \leq -\lambda(x)|V_{[\alpha]}f(x)|^2, \quad \forall x \in \mathbb{R}^N, \alpha \in \mathcal{A}_m. \tag{A.25}$$

However, as we show in the following lemma, for the case when $N = d = 1$ it is sufficient to only consider the vector field V_0 and (A.25) reduces to (1.17).

Lemma A.2.4. *Consider the one dimensional SDE* (4.3)*. If the UFG condition and the LOAC* (A.25) *hold then for all* $x \in \mathbb{R}$,

$$\mathrm{span}(V_1(x)) = \mathrm{span}(V_{[\alpha]}(x) : \alpha \in \mathcal{A}_m). \tag{A.26}$$

Proof of Lemma A.2.4. Fix some $x_0 \in \mathbb{R}$; if $V_1(x_0) \neq 0$ then we have $\mathrm{span}(V_1(x_0)) = \mathbb{R}$ and (A.26) follows immediately.

If $V_1(x_0) = 0$, we want to prove by induction that $V_{[\alpha]}(x_0) = 0$ for every $\alpha \in \mathcal{A}_m$. To this end, suppose $V_{[\alpha]}(x_0) = 0$ for some $\alpha \in \mathcal{A}_m$; then we may use Taylor's theorem to obtain the following expansions

$$V_{[\alpha]}(x) = V'_{[\alpha]}(x_0)(x - x_0) + O((x - x_0)^2)$$

$$V'_{[\alpha]}(x) = V'_{[\alpha]}(x_0) + V''_{[\alpha]}(x_0)(x - x_0) + O((x - x_0)^2)$$

$$V_0(x) = V_0(x_0) + V'_0(x_0)(x - x_0) + O((x - x_0)^2)$$

$$V'_0(x) = V'_0(x_0) + V''_0(x_0)(x - x_0) + O((x - x_0)^2)$$

$$V_1(x) = V'_1(x_0)(x - x_0) + O((x - x_0)^2).$$

Here $O((x-x_0)^n)$ denotes functions f such that for some neighbourhood of x_0, there is some constant $C > 0$ such that

$$|f(x)| \leq C|x - x_0|^n.$$

Substituting these expansions into the definition of $[V_{[\alpha]}, V_0]$ we have

$$[V_{[\alpha]}, V_0](x) = V_{[\alpha]}(x)V_0'(x) - V_0(x)V_{[\alpha]}'(x)$$
$$= -V_{[\alpha]}'(x_0)V_0(x_0) - V_{[\alpha]}''(x_0)V_0(x_0)(x - x_0) + O((x - x_0)^2).$$

Then, expanding the left hand side and right hand side of (A.25), we have

$$V_{[\alpha]}(x)[V_{[\alpha]}, V_0](x) = -|V_{[\alpha]}'(x_0)|^2 V_0(x_0)(x - x_0) - V_{[\alpha]}'(x_0)V_{[\alpha]}''(x_0)V_0(x_0)(x - x_0)^2$$
$$+ O((x - x_0)^3)$$
$$|V_{[\alpha]}(x)|^2 = V_{[\alpha]}'(x_0)^2(x - x_0)^2 + O((x - x_0)^3),$$

hence, by (1.17),

$$- |V_{[\alpha]}'(x_0)|^2 V_0(x_0)(x - x_0) - V_{[\alpha]}'(x_0)V_{[\alpha]}''(x_0)V_0(x_0)(x - x_0)^2$$
$$\leq -\lambda(x)V_{[\alpha]}'(x_0)^2(x - x_0)^2 + O((x - x_0)^3).$$

Rearranging the above gives

$$-V_0(x_0)V_{[\alpha]}'(x_0)^2 \frac{(x - x_0)}{|x - x_0|} \leq \left(V_{[\alpha]}'(x_0)V_0(x_0)V_{[\alpha]}''(x_0) - \lambda(x)V_{[\alpha]}'(x_0)^2\right) \frac{(x - x_0)^2}{|x - x_0|}$$
$$+ O((x - x_0)^2).$$

Suppose that $V_0(x_0)V_{[\alpha]}'(x_0) \neq 0$; then letting x tend to x_0, we obtain a contradiction. Therefore $V_0(x_0)V_{[\alpha]}'(x_0)$ must be equal to zero which implies that $[V_{[\alpha]}, V_0](x_0)$ is equal to zero as well. Moreover, since $V_1(x_0) = V_{[\alpha]}(x_0) = 0$, we also have $[V_{[\alpha]}, V_1](x_0) = 0$. Then by induction we have $V_{[\alpha]}(x_0) = 0$ for all $\alpha \in \mathcal{A}_m$. This concludes the proof. $\square$

Lemma A.2.5. *Consider the SDE* (4.19)*. If there exists $R > 0$ such that* $\operatorname{sign}(x)b(x) < 0$ *whenever $|x| \geq R$ then Hypothesis 4.2.7 holds with*

$$\Xi(x) = \alpha^2 + \alpha b(x) \tanh(\alpha x)\,,$$

where $\alpha > 0$ is any positive constant. Moreover, if b is unbounded (both above and below) then Hypothesis 4.2.3 also holds.

Proof of Lemma A.2.5. Let $u(x) = \cosh(\alpha x)$ and define $\theta : \mathbb{R} \to \mathbb{R}$ as in Note 4.2.4. Then Hypothesis 4.2.7 is satisfied with the functions

$$u_n(x) = u\left(n\theta\left(\frac{x}{n}\right)\right),$$

$$\Xi(x) = \alpha b(x)\tanh(\alpha x) + \alpha^2,$$

as we come to explain. By construction $u_n(x) \leq u(x)$ for all $x \in \mathbb{R}$ and $n \in \mathbb{N}$, so for each compact set $W \subseteq \mathbb{R}$

$$\sup_{x\in W}\sup_{n\in\mathbb{N}} u_n(x) \leq \sup_{x\in W} u(x) < \infty,$$

so Hypothesis 4.2.3 (1c). Now for fixed $x \in \mathbb{R}$ and $n > |x|$ we have $u_n(x) = u(x)$ and

$$\frac{\mathcal{L}u(x)}{u(x)} = \alpha b(x)\tanh(\alpha x) + \alpha^2 = \Xi(x),$$

therefore Hypothesis 4.2.3 (1d) is satisfied. Moreover we see that if $|x| \leq n$ then

$$\frac{\mathcal{L}u_n}{u_n} \leq \Xi(x) \leq A_1,$$

where A_1 is the maximum value of Ξ. Now if $|x| \geq 2n$ then $u_n(x)$ is constant and we have

$$\frac{\mathcal{L}u_n}{u_n} = 0.$$

However if $n \leq |x| \leq 2n$ then

$$\frac{\mathcal{L}u_n(x)}{u_n(x)} = \left(\alpha b(x)\theta'\left(\frac{x}{n}\right) + \alpha^2\frac{1}{n}\theta''\left(\frac{x}{n}\right)\right)\tanh\left(\alpha n\theta\left(\frac{x}{n}\right)\right) + \alpha^2\theta'\left(\frac{x}{n}\right).$$

We may assume that n is sufficiently large that $b(x)\tanh(\alpha n\theta(\frac{x}{n})) < 0$. Then since θ is increasing on $[1,2]$ and is an odd function we have that $b(x)\theta'(x/n)\tanh(\alpha n\theta(x/n)) < 0$ thus

$$\frac{\mathcal{L}u_n(x)}{u_n(x)} \leq \alpha^2\frac{1}{n}\theta''\left(\frac{x}{n}\right)\tanh\left(\alpha n\theta\left(\frac{x}{n}\right)\right) + \alpha^2\theta'\left(\frac{x}{n}\right) \leq \alpha^2 \sup_{y\in[1,2]}\left[|\theta''(y)| + \theta'(y)\right] =: A_2.$$

Therefore (4.32) holds with $A = \max\{A_1, A_2\}$.

Moreover, if b is unbounded (both above and below) and $\mathrm{sign}(x)b(x) < 0$ for x sufficiently large, we see that the set $\{x \in \mathbb{R} : \Xi(x) \geq \ell\}$ is compact for each $\ell \in \mathbb{R}$. $\square$

Proof of Lemma 4.1.10. We shall only prove (4.20) for $\partial_x^4 \mathcal{P}_t f(x)$ as the other estimates follow by a simpler version of the same argument. Recall J_t is defined as $J_t = \frac{\partial}{\partial x} X_t^{(x)}$ and we can similarly define higher order derivatives as $J_t^{(n)} = \frac{\partial^n}{\partial x^n} X_t^{(x)}$.

Using the chain rule and then taking expectations, similiarly to Lemma 4.1.3, we obtain

$$\partial_x^4 \mathcal{P}_t f(x) = \mathbb{E}\Big[f^{(4)}(X_t^{(x)}) J_t^4 + 6 f^{(3)}(X_t^{(x)}) J_t^2 J_t^{(2)} \tag{A.27}$$

$$+ 3 f''(X_t^{(x)}) (J_t^{(2)})^2 + 4 f''(X_t^{(x)}) J_t J_t^{(3)} + f'(X_t^{(x)}) J_t^{(4)} \Big] \tag{A.28}$$

$$\leq K \|f\|_{C_b^4(\mathbb{R})} \mathbb{E}\left[J_t^4 + J_t^2 |J_t^{(2)}| + (J_t^{(2)})^2 + J_t |J_t^{(3)}| + |J_t^{(4)}| \right]. \tag{A.29}$$

Now we proceed by estimating each of these terms in turn. Note that in the case at hand (4.7) simplifies to

$$dJ_t = b'(X_t^{(x)}) dt \tag{A.30}$$

which we can solve to find

$$J_t = \exp\left(\int_0^t b'(X_s^{(x)}) ds \right). \tag{A.31}$$

Differentiating (A.31) we obtain the following expressions for the higher order derivatives,

$$
\begin{aligned}
J_t^{(2)} &= \int_0^t b''(X_s^{(x)} J_s ds J_t \\
J_t^{(3)} &= \left(\int_0^t b^{(3)}(X_s^{(x)}) J_s^2 ds + \int_0^t b''(X_s^{(x)}) J_s^{(2)} ds \right) J_t \\
&= \int_0^t b''(X_s^{(x)}) J_s ds J_t^{(2)} \\
J_t^{(4)} &= \left(\int_0^t b^{(4)}(X_s^{(x)}) J_s^3 + 3b^{(3)}(X_s^{(x)}) J_s J_s^{(2)} + b''(X_s^{(x)}) J_s^{(3)} ds \right) J_t \\
&+ \left(\int_0^t 3b^{(3)}(X_s^{(x)}) J_s^2 + 2b''(X_s^{(x)}) J_s^{(2)} ds \right) J_t^{(2)} \\
&+ \int_0^t b''(X_s^{(x)}) J_s ds J_t^{(3)}.
\end{aligned}
$$

From here it is straight forward to see that the conclusion holds. □

A.2.3 Proofs of Section 5.1

We include here the proofs of Lemma A.2.6 and Lemma A.2.9, on which the proof of Theorem 5.1.5 hinges.

Lemma A.2.6. *Let Hypothesis 5.1.1* **[H.1]** *and* **[H.3]** *hold and assume the semigroup $\{Q_{s,t}\}_{0\le s\le t}$ admits an evolution system of measures $\{\nu_t\}_{t\ge 0}$; then, for each $g \in C_b(\mathbb{R}^n)$, $z \in \mathbb{R}^n$, and $s \ge 0$ we have*

$$\lim_{t\to\infty} \left| Q^{\zeta}_{s,t} g(z) - \int_{\mathbb{R}^n} g(y) d\nu_t \right| \to 0. \tag{A.32}$$

Proof of Lemma A.2.6. Fix $g \in C_b(\mathbb{R}^n)$ and let $f \in C_b(\mathbb{R}^{n+1})$ be a function that doesn't depend on the last variable and such that $f(z,\eta) = g(z)$ for every $\eta \in \mathbb{R}, z \in \mathbb{R}^n$; note that by (5.6) we have

$$Q^{\zeta}_{s,t} g(z) = \mathcal{P}_{t-s} f(z, \zeta^{\zeta}_s). \tag{A.33}$$

Now for every fixed $z, y \in \mathbb{R}^n$ we can write

$$|Q_{s,t}g(z) - Q_{s,t}g(y)| = |\mathcal{P}_{t-s}f(z,\zeta_s) - \mathcal{P}_{t-s}f(y,\zeta_s)|$$

By Note 5.1.2 we have that the hyperplane $S := \{x = (z,\zeta) \in \mathbb{R}^{n+1} : \zeta = \zeta_s\}$ is the orbit of the vector fields $V_{[\alpha]}$, $\alpha \in \mathcal{A}_m$. Since (z,ζ_s) and (y,ζ_s) belong to S we may take a piecewise integral curve connecting them. Without loss of generality we may take an integral curve $\gamma : [0,T] \to \mathbb{R}^{n+1}$ connecting (z,ζ_s) and (y,ζ_s), with $\dot{\gamma}_t = V_{[\alpha]}(\gamma_t)$. Clearly the time T will depend on z and y, i.e. $T = T_{z,y}$. Let K be a compact set such that $\gamma([0,T_{z,y}]) \subseteq K$; by Lemma A.1.15 applied to the function $h = \mathcal{P}_{t-s}f$, which is in $\mathcal{D}_V^{2,\infty}(\mathbb{R}^N)$ by Appendix A.1.4 **[F.3]**, we have

$$|\mathcal{P}_{t-s}f(z,\zeta_s) - \mathcal{P}_{t-s}f(y,\zeta_s)| \leq \int_0^{T_{z,y}} V_{[\alpha]}(\mathcal{P}_{t-s}f(\gamma_u))du.$$

Because we let $t \to \infty$, we can restrict to the case $t > s$. So fix $s_0 > 0$ such that $t - s > s_0$; by (2.9) we then have

$$|\mathcal{P}_{t-s}f(z,\zeta_s) - \mathcal{P}_{t-s}f(y,\zeta_s)| \leq c_{s_0,r} e^{-\lambda(t-s-s_0)} \|f\|_\infty T_{z,y}.$$

Letting t tend to ∞ and using (A.33) we obtain

$$\lim_{t\to\infty} |Q_{s,t}g(z) - Q_{s,t}g(y)| = 0. \tag{A.34}$$

The proof can now be concluded as follows: because $\{\nu_t\}_t$ is an evolution system of measures (see 5.12), we can write

$$\begin{aligned}\left|Q_{s,t}g(z) - \int_{\mathbb{R}^n} g(y)\nu_t(dy)\right| &= \left|Q_{s,t}g(z) - \int_{\mathbb{R}^n} Q_{s,t}g(y)\nu_s(dy)\right| \\ &\leq \int_{\mathbb{R}^n} |Q_{s,t}g(z) - Q_{s,t}g(y)|\,\nu_s(dy).\end{aligned}$$

Using (A.34) and the dominated convergence theorem (which is applicable by Appendix A.1.5 **[F.1]**) we may take the limit as t tends to ∞ and obtain (A.32). □

In order to prove Lemma A.2.9, which is the core of the last step of the proof of Theorem 5.1.5, we must first prove the following two results, Lemma A.2.7 and

Lemma A.2.8.

Lemma A.2.7. *Assume Hypothesis 5.1.1 holds. Then, for each $g \in C_b(\mathbb{R}^n)$ and $z \in \mathbb{R}^n$ we have*

$$Q_{s-t,s}g(z) \to \bar{Q}_t g(z)$$

as s tends to ∞ uniformly on compacts of $\mathbb{R}^n \times (0, \infty)$. That is, for every fixed $T > 0$ and $r > 0$ we have

$$\lim_{s\to\infty} \sup_{z\in B_r} \sup_{[1/T,T]} \left|Q_{s-t,s}g(z) - \bar{Q}_t g(z)\right| = 0.$$

Proof of Lemma A.2.7. Fix $g \in C_b(\mathbb{R}^n)$ and consider

$$v_s(z,t) = (Q^{\zeta}_{s-t,s}g)(z), \quad z \in \mathbb{R}^n, t > 0.$$

Like in the Proof of Lemma A.2.6, define $f \in C_b(\mathbb{R}^{n+1})$ by $f(z, \eta) = g(z)$ for all $z \in \mathbb{R}^n$ and $\eta \in \mathbb{R}$; then by (A.33) we have

$$v_s(z,t) = \mathcal{P}_t f(z, \zeta_{s-t}). \tag{A.35}$$

Since $(\mathcal{P}_t f)(x)$ is a continuous function, we may take the limit as s tends to infinity to obtain

$$\lim_{s\to\infty} v_s = \mathcal{P}_t f(z, \bar{\zeta}) = \bar{Q}_t g(z) =: v(z,t).$$

We now wish to show that the above limit is uniform on compact subsets of $\mathbb{R}^n \times (0, \infty)$; that is, we wish to show that

$$\lim_{s\to\infty} \sup_{t\in[\frac{1}{T},T]} \sup_{z\in B_R} |v_s(z,t) - v(z,t)| = 0, \quad \text{for every fixed } R > 0, T > 0.$$

To show this fact we shall use the Ascoli-Arzela Theorem. Indeed, assuming for the moment that we can apply such a theorem, then we can find a subsequence s_k such that v_{s_k} converges uniformly on $B_R \times [1/T, T]$. Since v_{s_k} converges pointwise to v we have that the limit is independent of the choice of sequence hence v_s converges uniformly in $B_R \times [1/T, T]$ to v, as s tends to ∞. So, if we show that the derivatives of v_s are bounded on $B_R \times [\frac{1}{T}, T]$, uniformly in s, then we may apply Arzela-Ascoli

Theorem and the proof is concluded by the above line of reasoning. By Lemma A.1.16 the function v_s is smooth in $(z,t) \in \mathbb{R}^n \times \mathbb{R}$ and satisfies (A.6).

For any point $(z,t) \in \mathbb{R}^n \times (0,\infty)$ there exist an open neighbourhood of (z,t) and smooth functions $\varphi_{i,\alpha}$ such that, for any $i \in \{1,\ldots,n\}$, the derivative $\partial_i \equiv \partial_{z^i}$ can be expressed as

$$\partial_i = \sum_{\alpha \in \mathcal{A}_m} \varphi_{i,\alpha} V_{[\alpha]}.$$

Therefore, to show that the derivatives $\partial_i v_s$ are bounded on $B_R \times [0,T]$ it is sufficient to show that $V_{[\alpha]} v_s$ is bounded in $B_R \times [1/T, T]$. This follows from the estimates recalled in Appendix A.1.4 **[F.2]**. In particular, there exists constants $C(R), \omega(R) > 0$ such that

$$\begin{aligned}
\sup_{z \in B_R, t \in [1/T,T]} \left| V_{[\alpha]} v_s(z,t) \right| &= \sup_{z \in B_R, t \in [1/T,T]} \left| V_{[\alpha]} \mathcal{P}_t f(z, \zeta_{s-t}) \right| \\
&\leq \sup_{t \in [1/T,T]} \sup_{x \in B_R \times \{\zeta_t : t \geq -T\}} \left| V_{[\alpha]} \mathcal{P}_t f(x) \right| \\
&\leq C(R) |T|^{\frac{\|\alpha\|}{2}} e^{\omega(R)T} \|f\|_\infty.
\end{aligned}$$

Here we have used that ζ_t is convergent and hence $B_R \times \{\zeta_t : t \geq -T\}$ is a compact subset of $\mathbb{R}^{n+1}$. Similarly we may bound the second order derivatives $V_i^2 v_s$, and using (A.6) we obtain a bound for the derivative with respect to t, which is independent of s. □

Using the tightness of the family $\{\nu_t\}_{t \geq 0}$, there exists a divergent sequence t_ℓ such that $\nu_{t_\ell - k}$ converges weakly to some measure μ_k as ℓ tends to ∞, for each $k \in \mathbb{N}$. (We emphasise that, by a diagonal argument, the sequence t_ℓ can be chosen to be independent of k). Moreover, $\{\mu_k\}_{k \in \mathbb{N}}$ is tight since $\{\nu_t\}_{t \geq 0}$ is tight (see [62, Step 2 in the proof of Theorem 6.2]).

Lemma A.2.8. *Assume Hypothesis 5.1.1 holds and construct $\{\mu_k\}_{k \in \mathbb{N}}$ as above. Then,*

$$\int_{\mathbb{R}^n} \bar{Q}_k g(z) \mu_k(dz) = \int g(z) \mu_0(dz),$$

for any $g \in C_b(\mathbb{R}^n)$ and every $k \in \mathbb{N}$.

Proof of Lemma A.2.8. We will consider the integral $\int_{\mathbb{R}^n} Q_{t_\ell - k, t_\ell} g(z) \nu_{t_\ell - k}(dz)$ and show the following:

$$\int_{\mathbb{R}^n} g(z)\mu_0(dz) = \lim_{\ell\to\infty}\int_{\mathbb{R}^n} Q_{t_\ell-k,t_\ell}g(z)\nu_{t_\ell-k}(dz) = \int_{\mathbb{R}^n} \bar{Q}_k g(z)\mu_k(dz)\,, \quad \text{for every } k\in\mathbb{N}. \tag{A.36}$$

Let us start with showing the first equality in (A.36). Because $\{\nu_t\}_{t\geq 0}$ is an evolution system of measures (and taking ℓ sufficiently large that $t_\ell > k$), we have

$$\int_{\mathbb{R}^n} Q_{t_\ell-k,t_\ell}g(z)\nu_{t_\ell-k}(dz) = \int g(z)\nu_{t_\ell}(dz).$$

The above, combined with the fact that ν_{t_ℓ} converges weakly to μ_0, gives the first identity in (A.36). To prove the second equality in (A.36), observe the following:

$$\begin{aligned}\int_{\mathbb{R}^n} Q_{t_\ell-k,t_\ell}g(z)\nu_{t_\ell-k}(dz) - \int_{\mathbb{R}^n} \bar{Q}_k g(z)\mu_k(dz) =& \int_{\mathbb{R}^n}\left(Q_{t_\ell-k,t_\ell}g(z) - \bar{Q}_k g(z)\right)\nu_{t_\ell-k}(dz)\\ &+\int_{\mathbb{R}^n} \bar{Q}_k g(z)\nu_{t_\ell-k}(dz) - \int_{\mathbb{R}^n} \bar{Q}_k g(z)\mu_k(dz)\\ =& I_{1,\ell} + I_{2,\ell},\end{aligned}$$

having set

$$\begin{aligned} I_{1,\ell} &:= \int_{\mathbb{R}^n}\left(Q_{t_\ell-k,t_\ell}g(z) - \bar{Q}_k g(z)\right)\nu_{t_\ell-k}(dz),\\ I_{2,\ell} &:= \int_{\mathbb{R}^n} \bar{Q}_k g(z)\nu_{t_\ell-k}(dz) - \int_{\mathbb{R}^n} \bar{Q}_k g(z)\mu_k(dz).\end{aligned}$$

Now $I_{2,\ell}$ converges to 0 as $\ell\to\infty$ since $\nu_{t_\ell-k}$ converges weakly to μ_k, by definition of μ_k. To see that $I_{1,\ell}$ vanishes when ℓ tends to ∞ fix $\varepsilon>0$ and take a ball B_r such that $\nu_{t_\ell-k}(B_r)\geq 1-\varepsilon$ for all ℓ with $t_\ell > k$. This is possible since the family $\{\nu_t : t\geq 0\}$ is tight. By Lemma A.2.7 we know that $Q_{t_\ell-k,t_\ell}g(z)$ converges uniformly on compacts to $\bar{Q}_k g(z)$; hence, if ℓ is sufficiently large, we have

$$\sup_{z\in B_r}|Q_{t_\ell-k,t_\ell}g(z) - \bar{Q}_k g(z)| \leq \varepsilon.$$

We can therefore derive the following estimate in $I_{1,\ell}$:

$$
\begin{aligned}
I_{1,\ell} = \int_{\mathbb{R}^n} \left(Q_{t_\ell - k, t_\ell} g(z) - \bar{Q}_k g(z)\right) \nu_{t_\ell - k}(dz) = & \int_{B_r} \left(Q_{t_\ell - k, t_\ell} g(z) - \bar{Q}_k g(z)\right) \nu_{t_\ell - k}(dz) \\
& + \int_{\mathbb{R}^n \setminus B_r} \left(Q_{t_\ell - k, t_\ell} g(z) - \bar{Q}_k g(z)\right) \nu_{t_\ell - k}(dz) \\
& \leq \varepsilon + 2\|g\|_\infty \varepsilon.
\end{aligned}
$$

As ε is arbitrary we have that $I_{1,\ell}$ converges to 0 as ℓ tends to ∞, and the claim follows. □

Lemma A.2.9. *Assume Hypothesis 5.1.1 holds and, as described before the statement of Lemma A.2.8, let μ_0 be the weak limit of the sequence ν_{t_ℓ}. Then $\mu_0 = \bar{\mu}$.*

Proof of Lemma A.2.9. Take $g \in C_b(\mathbb{R}^n)$. By Lemma 5.1.4 we know that $\bar{Q}_k g(z) \to \mu(g)$ as k tends to ∞ for each $z \in \mathbb{R}^n$ and $g \in C_b(\mathbb{R}^n)$. By an argument analogous to the one used in the proof of Lemma A.2.7 we have that $\bar{Q}_k g(z)$ converges to $\mu(g)$ locally uniformly for $z \in \mathbb{R}^n$.

Now fix $\varepsilon > 0$; since $\{\mu_k\}_k$ is a tight sequence, we may take $B_r \subseteq \mathbb{R}^n$ such that $\mu_k(B_r) \geq 1 - \varepsilon$ for all $t \geq 0$. Moreover, for k sufficiently large we have

$$
\sup_{x \in B_r} |\bar{Q}_k g(z) - \bar{\mu}(g)| \leq \varepsilon.
$$

Then

$$
\begin{aligned}
\int_{\mathbb{R}^n} \bar{Q}_k g(z) \mu_k(dz) - \bar{\mu}(g) &= \int_{B_r} \left[\bar{Q}_k g(z) - \bar{\mu}(g)\right] \mu_k(dz) + \int_{\mathbb{R}^n \setminus B_r} \left[\bar{Q}_k g(z) - \bar{\mu}(g)\right] \mu_k(dz) \\
&\leq \varepsilon + 2\|g\|_\infty \varepsilon.
\end{aligned}
$$

Since ε is arbitrary we deduce

$$
\int_{\mathbb{R}^n} \bar{Q}_k g(z) \mu_k(dz) \to \bar{\mu}(g), \quad \text{as } k \to \infty.
$$

However by Lemma A.2.8 we also have

$$\int_{\mathbb{R}^n} g(z)\mu_0(dz) = \int_{\mathbb{R}^n} \bar{Q}_k g(z)\mu_k(dz), \qquad \text{for every } k \in \mathbb{N}.$$

Therefore $\bar{\mu} = \mu_0$. □

A.2.4 Proofs of Section 5.2

Proof of Proposition 5.3.1. First note to prove (5.21) it is sufficient to show that for any $\alpha \in \mathcal{A}_m$,

$$[\mathrm{Ad}_{tV_0^{(\perp)}} V_{[\alpha]}, \partial_t + \mathcal{V}_{0,t}] = \mathrm{Ad}_{tV_0^{(\perp)}}[V_{[\alpha]}, V_0], \tag{A.37}$$

$$[\mathrm{Ad}_{tV_0^{(\perp)}} V_{[\alpha]}, \mathcal{V}_{i,t}] = \mathrm{Ad}_{tV_0^{(\perp)}}[V_{[\alpha]}, V_i], \qquad \text{for any } i \in \{1, \dots, d\}. \tag{A.38}$$

By [69, Proposition 8.30], for any two vector fields U, V, W we have

$$[\mathrm{Ad}_W U, \mathrm{Ad}_W V] = \mathrm{Ad}_W [U, V]. \tag{A.39}$$

Therefore setting $W = tV_0^{(\perp)}$, $U = V_{[\alpha]}$ and $V = V_i$ for any $i \in \{1, \dots, d\}$ we have that (A.38) holds. To prove (A.37) note that

$$[\mathrm{Ad}_{tV_0^{(\perp)}} V_{[\alpha]}, \partial_t] = -\partial_t \mathrm{Ad}_{tV_0^{(\perp)}} V_{[\alpha]} = -\mathrm{Ad}_{tV_0^{(\perp)}}[V_0^{(\perp)}, V_{[\alpha]}] = \mathrm{Ad}_{tV_0^{(\perp)}}[V_{[\alpha]}, V_0^{(\perp)}]. \tag{A.40}$$

In the above we have used [57, Lemma 4.4.2] which states that $\partial_t \mathrm{Ad}_{tV_0^{(\perp)}} V_{[\alpha]} = \mathrm{Ad}_{tV_0^{(\perp)}}[V_0^{(\perp)}, V_{[\alpha]}]$. Using (A.39) and (A.40) we have

$$\begin{aligned}
[\mathrm{Ad}_{tV_0^{(\perp)}} V_{[\alpha]}, \partial_t + \mathrm{Ad}_{tV_0^{(\perp)}} V_0^{(\hat{\Delta})}] &= \mathrm{Ad}_{tV_0^{(\perp)}}[V_{[\alpha]}, V_0^{(\perp)}] + \mathrm{Ad}_{tV_0^{(\perp)}}[V_{[\alpha]}, V_0^{(\hat{\Delta})}] \\
&= \mathrm{Ad}_{tV_0^{(\perp)}}[V_{[\alpha]}, V_0].
\end{aligned}$$

Therefore (A.37) holds. Hence we have that (5.21) holds.

It remains to show that for any $\alpha \in \mathcal{A}$ there exists smooth and bounded functions

$\tilde{\varphi}_{\alpha,\beta} \in C_V^\infty(\mathbb{R}^N \times \mathbb{R})$ such that

$$\mathcal{V}_{[\alpha],t}(x) = \sum_{\beta \in \mathcal{A}_m} \tilde{\varphi}_{\alpha,\beta}(x,t)\mathcal{V}_{[\beta],t}(x).$$

Since the vector fields $\{V_0, V_1, \ldots, V_d\}$ satisfy the UFG condition (**UFG**), applying the operator $\mathrm{Ad}_{tV_0^{(\perp)}}$ to (**UFG**) we have

$$\mathrm{Ad}_{tV_0^{(\perp)}} V_{[\alpha]}(x) = \sum_{\beta \in \mathcal{A}_m} \mathrm{Ad}_{tV_0^{(\perp)}}(\varphi_{\alpha,\beta} V_{[\beta]})(x) = \sum_{\beta \in \mathcal{A}_m} \varphi_{\alpha,\beta}(e^{tV_0^{(\perp)}}(x))\mathrm{Ad}_{tV_0^{(\perp)}} V_{[\beta]}(x).$$

Therefore if we show that the functions $\tilde{\varphi}(x,t) := \varphi_{\alpha,\beta}(e^{tV_0^{(\perp)}}(x))$ are smooth, bounded and belong to the sets $C_V^\infty(\mathbb{R}^N \times \mathbb{R})$ then we have that the vector fields $\{\partial_t + \mathcal{V}_{[0],t}, \mathcal{V}_{[1],t}, \ldots, \mathcal{V}_{[d],t}\}$ satisfy the UFG condition when viewed as vector fields in both the time variable t and spatial variables z. Since $\varphi_{\alpha,\beta}$ is smooth and bounded, and $V_0^{(\perp)}$ is smooth we have that $\varphi_{\alpha,\beta} \circ e^{tV_0^{(\perp)}}$ is smooth and bounded, it remains to show that for any $k \in \mathbb{N}$ and $\gamma_1, \ldots, \gamma_k \in \mathcal{A}$ we have

$$\sup_{x \in \mathbb{R}^N, t \in \mathbb{R}} \left| (\mathcal{V}_{[\gamma_1],t}) \ldots (\mathcal{V}_{[\gamma_k],t})(\varphi_{\alpha,\beta} \circ e^{tV_0^{(\perp)}}) \right| < \infty$$

By [69, Proposition 8.30] for every smooth function f we have for all $\alpha \in \mathcal{A}$, $z \in S_{x_0}$, $s \in \mathbb{R}$ that

$$\mathcal{V}_{[\alpha],s}(f \circ e^{sV_0^{(\perp)}})(z) = (V_{[\alpha]} f)(e^{sV_0^{(\perp)}}(z)). \tag{A.41}$$

Therefore

$$\begin{aligned}\sup_{x \in \mathbb{R}^N, s \in \mathbb{R}} \left| (\mathcal{V}_{[\gamma_1],t}) \ldots (\mathcal{V}_{[\gamma_k,t]})(\varphi_{\alpha,\beta} \circ e^{tV_0^{(\perp)}}) \right| &= \sup_{x \in \mathbb{R}^N, t \in \mathbb{R}} \left| (V_{[\gamma_1]} \ldots V_{[\gamma_k]} \varphi_{\alpha,\beta})(e^{tV_0^{(\perp)}}(x)) \right| \\ &= \sup_{y \in \mathbb{R}^N} \left| (V_{[\gamma(1)]} \ldots V_{[\gamma(k)]} \varphi_{\alpha,\beta})(y) \right| < \infty.\end{aligned}$$

Therefore the UFG condition is satisfied. □

Proof of Proposition 5.3.2. For $f \in C_V^\infty(\mathbb{R}^N)$ we have that $\mathcal{P}_t f$ is smooth (in every direction, see [**F.3**]) so by (5.20) we obtain

$$\mathcal{V}_{[\alpha],s}\mathcal{Q}_{s,t}f(z) = \mathcal{V}_{[\alpha],s}\left(\mathcal{P}_{t-s}(f\circ e^{-tV_0^{(\perp)}}))(e^{sV_0^{(\perp)}})\right)(z)$$

$$\overset{(A.41)}{=} V_{[\alpha]}\left(\mathcal{P}_{t-s}(f\circ e^{-tV_0^{(\perp)}})\right)(e^{sV_0^{(\perp)}}(z)). \tag{A.42}$$

By differentiating (5.20) with respect to s we have, with a calculation analogous to the one in Lemma A.1.16,

$$\partial_s\mathcal{Q}_{s,t}f(z) = -\mathcal{L}\mathcal{P}_{t-s}(f\circ e^{-tV_0^{(\perp)}})(e^{sV_0^{(\perp)}}(z)) + V_0^{(\perp)}\mathcal{P}_{t-s}(f\circ e^{-tV_0^{(\perp)}})(e^{sV_0^{(\perp)}}(z))$$

$$\overset{(1.12)}{=} -V_0^{(\hat{\Delta})}\mathcal{P}_{t-s}(f\circ e^{-tV_0^{(\perp)}})(e^{sV_0^{(\perp)}}(z)) - \sum_{i=1}^d V_i^2\mathcal{P}_{t-s}(f\circ e^{-tV_0^{(\perp)}})(e^{sV_0^{(\perp)}}(z))$$

$$= -\mathcal{V}_{0,s}\mathcal{Q}_{s,t}f(z) - \sum_{i=1}^d \mathcal{V}_{i,s}\mathcal{Q}_{s,t}f(z) = -\mathcal{L}_s\mathcal{Q}_{s,t}f(z).$$

Now by a density argument analogous to the one in the proof of Lemma A.1.16 we obtain the result for $f \in C_b(\mathbb{R}^N)$. To prove the result for $g \in C_b(\overline{S}_{x_0})$ we may apply the Tietze Extension Theorem, see [70, Chapter 2 Theorem 5.4], to extend g to a function $f \in C_b(\mathbb{R}^N)$ such that $f = g$ on $\overline{S}_{x_0}$ (this is where we need g to be continuous up to and including the closure of S_{x_0}, as functions that are continuous on open sets don't necessarily admit a continuous extension to the whole $\mathbb{R}^N$, i.e. Tietze Extension Theorem would not apply). Since $\mathcal{Z}_t$ takes values in S_{x_0} for every $t \geq 0$ we have that $\mathcal{Q}_{s,t}g(z) = \mathcal{Q}_{s,t}f(z)$ for any $z \in \overline{S}_{x_0}$, hence the claim follows. □

Proof of Proposition 5.3.5. By Hypothesis 5.3.3 [**A.3**], the family of measures $\{p_t^x\}_{t\geq 0}$ is tight and hence, by Prokhorov's Theorem, there exist a measure $\overline{\mu}^S$ and a diverging sequence $\{t_k\}_k$ such that $p_{t_k}^x$ converges weakly to $\overline{\mu}^S$ as $t_k \nearrow \infty$. Note that in general the sequence t_k and the measure $\overline{\mu}^S$ may depend on the choice of $x \in S_{x_0}$; however, by Lemma 3.2.9, $p_{t_k}^x(\cdot) = (\mathcal{P}_t\mathbb{1}_{\{\cdot\}})(x)$ converges weakly to $\overline{\mu}^S$ for any choice of $x \in S$. We now show that such a convergence is also independent of the choice of divergent sequence. Let s_k be a sequence such that $s_k \nearrow \infty$ and fix $f \in C_b(\mathbb{R}^N)$

and $x \in S$; then

$$\mathcal{P}_{t_k} f(x) - \mathcal{P}_{s_k} f(x) = \int_{s_k}^{t_k} \partial_t \mathcal{P}_t f(x) dt = \int_{s_k}^{t_k} \mathcal{L}\mathcal{P}_t f(x) dt.$$

By (2.9) (and (2.11)) there exists a constant $C = C(t_0, x) > 0$ such that for all $t > t_0$ we have

$$\left| V_0^{(\hat{\Delta})} \mathcal{P}_t f(x) \right| \leq C(t_0, x) \|f\|_\infty e^{-\lambda t}, \quad \left| V_i^2 \mathcal{P}_t f(x) \right| \leq C(t_0, x) \|f\|_\infty e^{-\lambda t}.$$

Using that $V_0^{(\perp)} = 0$ we have that there exists a constant $C = C(t_0, x) > 0$ such that

$$|\mathcal{L}\mathcal{P}_t f(x)| \leq C(t_0, x) \|f\|_\infty e^{-\lambda t}, \qquad \text{for all } t > t_0.$$

Therefore

$$|\mathcal{P}_{t_k} f(x) - \mathcal{P}_{s_k} f(x)| \leq C(x, t_0) \|f\|_\infty \left| \int_{s_k}^{t_k} e^{-\lambda t} dt \right| = \frac{C(x, t_0) \|f\|_\infty}{\lambda} \left| e^{-\lambda t_k} - e^{-\lambda s_k} \right|.$$

Letting k tend to ∞ we have that $|\mathcal{P}_{t_k} f(x) - \mathcal{P}_{s_k} f(x)|$ vanishes in the limit and hence $\mathcal{P}_{s_k} f(x)$ converges to $\overline{\mu}^S(f)$. Therefore $\mathcal{P}_t f(x)$ converges to $\overline{\mu}^S(f)$ as t tends to ∞.

To show that $\overline{\mu}^S$ is an invariant measure take an arbitrary $s > 0$ and $f \in C_b(\mathbb{R}^N)$; then

$$\overline{\mu}^S(\mathcal{P}_s(f)) = \lim_{t \to \infty} \mathcal{P}_t \mathcal{P}_s f(x) = \lim_{t \to \infty} \mathcal{P}_{t+s} f(x) = \lim_{t \to \infty} \mathcal{P}_t f(x) = \overline{\mu}^S(f), \quad \text{for every } s \geq 0\,.$$

Hence $\overline{\mu}^S$ is an invariant measure. To show that the convergence is uniform on compact subsets of S we apply Arzela-Ascoli. Indeed fix a compact set $K \subseteq S$ then it is sufficient to show that $\mathcal{P}_t f(x)$ has bounded derivatives uniformly in t on K. However $x \mapsto \mathcal{P}_t f(x)$ is differentiable in the directions $V_{[\alpha]}$ for all $\alpha \in \mathcal{A}$ which span the tangent space of S and, by the Obtuse Angle Condition, Assumption [**A.5**], we have for all $t > t_0$ that (2.9) holds. Hence we have that $\mathcal{P}_t f(x)$ converges to $\overline{\mu}^S(f)$ uniformly on compact subsets of S. Note that since (5.24) holds for all $f \in C_b(\mathbb{R}^N)$, there is at most one measure satisfying (5.24). $\square$

We now move on to prove Lemma A.2.12 and Lemma A.2.13, which are the backbone of the proof of Theorem 5.3.7). Throughout this section, for any $f \in C_b(\overline{S}_{\overline{x}})$, we let

$$\hat{f} = f \circ W^{\infty}. \tag{A.43}$$

In order to prove Lemma A.2.12 we first state and prove the following two results.

Lemma A.2.10. *Let Hypothesis 5.3.3* **[A.1]**, **[A.2]**, **[A.7]** *and* **[A.6]** *hold. For any fixed $f \in C_b(\overline{S}_{\overline{x}})$, define $\hat{f}$ as in* (A.43)*. Then for any compact $K \subseteq S_{x_0}$ and $T > 0$ we have*

$$\mathcal{Q}_{t-s,t}\hat{f}(z) \to \mathcal{P}_s\hat{f}(\lim_{\tau\to\infty} e^{\tau V_0^{(\perp)}}(z)) \tag{A.44}$$

uniformly for $s \in [1/T, T]$ and $z \in K$, whenever $\lim_{\tau\to\infty} e^{\tau V_0^{(\perp)}}(z)$ exists for all $z \in K$.

Proof of Lemma A.2.10. The proof is analogous to the proof of Lemma A.2.7, so we only sketch it and point out the main differences. Note that $\mathcal{P}_t\hat{f}$ is continuous, using (5.20) we have that (A.44) holds pointwise. To obtain convergence uniform on compact subsets of $S_{x_0} \times (0, \infty)$, we use Arzela-Ascoli and the following estimate.

Fix a compact set $K \subseteq S_{x_0}$ and $T > 0$ then using (A.42) and the short time estimates from [50, Corollary 3.13] (which have been recalled in **[F.2]**), there exists some constant $C(K, T)$ such that the following holds:

$$\begin{aligned}
\sup_{z\in K, s\in[1/T,T]} |\mathcal{V}_{[\alpha],t-s}(\mathcal{Q}_{t-s,t}\hat{f})(z)| &= \sup_{z\in K, s\in[1/T,T]} |V_{[\alpha]}(\mathcal{P}_s(\hat{f}\circ e^{-tV_0^{(\perp)}}))(e^{(t-s)V_0^{(\perp)}}(z))| \\
&= \sup_{s\in[1/T,T], x\in e^{(t-s)V_0^{(\perp)}}(K)} |V_{[\alpha]}(\mathcal{P}_s(\hat{f}\circ e^{-tV_0^{(\perp)}}))(x)| \\
&= \sup_{s\in[1/T,T], x\in K'} |V_{[\alpha]}(\mathcal{P}_s(\hat{f}\circ e^{-tV_0^{(\perp)}}))(x)| \\
&\leq C(K', T)\|\hat{f}\|_{\infty};
\end{aligned}$$

here K' is defined as $K' = \bigcup_{\tau\geq -T} e^{\tau V_0^{(\perp)}}(K)$. The set K' is compact under our assumptions, as for each τ the diffeomorphism $x \to e^{\tau V_0^{(\perp)}}(x)$ is a continuous function, the curve $e^{\tau V_0^{(\perp)}}x$ is convergent and the map W^{∞} is assumed continuous. □

Define μ_k to be the probability measure such that $\nu_{t_\ell - k}$ converges weakly to μ_k, this measure is constructed analogously to the comment above Lemma A.2.8.

Lemma A.2.11. *Let Hypothesis 5.3.3* [**A.1**] *and* [**A.7**] *hold, and assume the semigroup* $\{\mathcal{Q}_{s,t}\}_{s\leq t}$ *admits a tight evolution system of measures* $\{\nu_t\}_{0\leq t}$ *supported on* S_{x_0}*. Then*

$$\int_{\overline{S}_{\overline{x}}} \mathcal{P}_k\hat{f}(x)\,(\mu_k\circ(W^\infty)^{-1})(dx) = \int_{\overline{S}_{\overline{x}}} f(x)\,(\mu_0\circ(W^\infty)^{-1})(dx),$$

for any $f\in C_b(\overline{S}_{\overline{x}})$ *and* $\hat{f}$ *defined as in* (A.43).

Proof of Lemma A.2.11. This proof is completely analogous to the proof of Lemma A.2.8, so we only point out the main differences. It suffices to prove the following two expressions

$$\int_{\overline{S}_{x_0}} \hat{f}(z)\mu_0(dz) = \lim_{\ell\to\infty}\int_{\overline{S}_{x_0}} \mathcal{Q}_{t_\ell-k,t_\ell}\hat{f}(z)\nu_{t_\ell-k}(dz) = \int_{\overline{S}_{x_0}} \mathcal{P}_k\hat{f}(W^\infty(z))\mu_k(dz), \tag{A.45}$$

compare to (A.36) for comparison. Let us start with the first equality in (A.45). Since $\{\nu_t\}_{t\geq 0}$ is an evolution system of measures we have

$$\int_{\overline{S}_{x_0}} \mathcal{Q}_{t_\ell-k,t_\ell}\hat{f}(z)\nu_{t_\ell-k}(dz) = \int_{\overline{S}_{x_0}} \hat{f}(z)\nu_{t_\ell}(dz).$$

Since ν_{t_ℓ} converges weakly to μ_0 and W^∞ is a continuous map from $\overline{S}_{x_0}$ to $\mathbb{R}^N$, by the continuous mapping theorem we have that $\nu_{t_\ell}\circ(W^\infty)^{-1}$ converges weakly to $\mu_0\circ(W^\infty)^{-1}$ and hence we obtain (A.45). To prove the second equality in (A.45) like in the proof of Lemma A.2.8 we write

$$\begin{aligned}
\int_{\overline{S}_{x_0}} \mathcal{Q}_{t_\ell-k,t_\ell}\hat{f}(z)\nu_{t_\ell-k}(dz) - &\int_{\overline{S}_{x_0}} \mathcal{P}_k\hat{f}(W^\infty(z))\mu_k(dz)\\
&= \int_{\overline{S}_{x_0}} \Big(\mathcal{Q}_{t_\ell-k,t_\ell}\hat{f}(x) - \mathcal{P}_k\hat{f}(W^\infty(x))\Big)\,\nu_{t_\ell-k}(dz)\\
&\quad + \int_{\overline{S}_{x_0}} \mathcal{P}_t\hat{f}(W^\infty(z))\nu_{t_\ell-k}(dz) - \int_{\overline{S}_{x_0}} \mathcal{P}_t\hat{f}(W^\infty(z))\mu_k(dz)\\
&= I_{1,\ell} + I_{2,\ell},
\end{aligned}$$

having set

$$I_{1,\ell} := \int_{\overline{S}_{x_0}} \left(\mathcal{Q}_{t_\ell - k, t_\ell} \hat{f}(z) - \mathcal{P}_k \hat{f}(W^\infty(z)) \right) \nu_{t_\ell - k}(dz)$$

$$I_{2,\ell} := \int_{\overline{S}_{x_0}} \mathcal{P}_k \hat{f}(W^\infty(z)) \nu_{t_\ell - k}(dz) - \int_{\overline{S}_{x_0}} \mathcal{P}_k \hat{f}(W^\infty(z)) \mu_k(dz).$$

Observe that on the image of W^∞ we have $V_0^{(\perp)} = 0$, by Hypothesis 5.3.3 [**A.7**], and hence $\mathcal{P}_k \hat{f}(W^\infty(z)) = \mathcal{P}_k f(W^\infty(z))$ therefore we can rewrite $I_{2,\ell}$ as

$$I_{2,\ell} = \int_{\overline{S}_{\overline{x}}} \mathcal{P}_k f(z) \left(\nu_{t_\ell - k} \circ (W^\infty)^{-1} \right) (dz) - \int_{\overline{S}_{\overline{x}}} \mathcal{P}_k \hat{f}(W^\infty(z)) \left(\mu_k \circ (W^\infty)^{-1} \right) (dz).$$

Now $I_{2,\ell}$ converges to 0 as $\ell \to \infty$ since $\nu_{t_\ell - k} \circ (W^\infty)^{-1}$ converges weakly to $\mu_k \circ (W^\infty)^{-1}$. The term $I_{1,\ell}$ can be studied analogously to what we have done in the proof of Lemma A.2.8, up to modifications in the same spirit of those made so far, so we omit the details.

□

Lemma A.2.12. *Suppose Hypothesis 5.3.3 holds. Let $\overline{\mu}^{S_{\overline{x}}}$ be defined as in the comment above Theorem 5.3.7. Then $\mu_0 \circ (W^\infty)^{-1} = \overline{\mu}^{S_{\overline{x}}}$.*

Proof of Lemma A.2.12. This proof is completely analogous to the proof of Lemma A.2.9 so we omit the details. □

Lemma A.2.13. *Assume Hypothesis 5.3.3 holds, let x_0 be an arbitrary point in $\mathcal{I}_0(\overline{x})$ and let $\{\nu_t\}, \mu_0$ be constructed as in the proof of Theorem 5.3.7. Let $\{t_\ell\}$ be a divergent sequence such that $p_{t_\ell}^{x_0}$ converges weakly to some probability measure ν^{x_0}. Then $\nu^{x_0}|_{\overline{S}_{\overline{x}}} = \mu_0 \circ (W^\infty)^{-1}$.*

Proof of Lemma A.2.13. First we note that ν^{x_0} and $\mu_0 \circ (W^\infty)^{-1}$ are both supported on $\overline{S}_{\overline{x}}$. Indeed for the measure $\mu_0 \circ (W^\infty)^{-1}$ this follows from Lemma A.2.12. It is sufficient to show that given a function $f \in C_b(\mathbb{R}^N)$ such that $f(x) = f(y)$ whenever $x \in \overline{S}_y$ then $\nu^{x_0}(f) = f(W^\infty(x_0))$. Let f be such a function then by Proposition 3.2.3 we have

$$\mathcal{P}_{t_\ell} f(x_0) = \mathbb{E}_{x_0}[f(X_{t_\ell})] = \mathbb{E}_{x_0}[f(e^{t_\ell V_0^{(\perp)}}(x_0))] = f(e^{t_\ell V_0^{(\perp)}}(x_0)).$$

Now letting ℓ tend to ∞ we have $\nu^{x_0}(f) = f(W^\infty(x_0))$ and hence ν^{x_0} must be supported on $\overline{S}_{\overline{x}}$. We now show that ν^{x_0} and $\mu_0 \circ (W^\infty)^{-1}$ coincide on $\overline{S}_{\overline{x}}$. Take a function $f \in C_b(\mathbb{R}^N)$ and let $\hat{f} = f \circ W^\infty$ then it is sufficient to show that $\nu^{x_0}(\hat{f}) = \mu_0(\hat{f})$. This follows from

$$\nu^{x_0}(\hat{f}) = \lim_{\ell\to\infty} \mathcal{P}_{t_\ell}\hat{f}(x_0) = \lim_{\ell\to\infty} \mathcal{Q}_{0,t_\ell}(\hat{f} \circ e^{t_\ell V_0^{(\perp)}})(x_0)$$

$$= \lim_{\ell\to\infty} \mathcal{Q}_{0,t_\ell}\hat{f}(x_0) = \mu_0(\hat{f}).$$

□

A.2.5 Proofs of Chapter 6

Proof of Lemma 6.1.1. Recall that $V_i = (U_i, 0)$ for $i = 1, \ldots, d$ and $V_0 = (U_0, W_0)$ where W_0 is independent of the variable z. By (6.1) we have

$$D_r^j X_t^{n+1} = D_r^j \zeta_t = \int_r^t \partial_{x^{n+1}} W_0(\zeta_s) D_r^j(\zeta_s) ds.$$

The only solution to this differential equation is $D_r^j \zeta_t = 0$. Therefore we have $\mathscr{M}_t^{i,n+1} = 0$ and $\mathscr{M}_t^{n+1,j} = 0$ for any $i, j = 1, \ldots, n+1$. Hence $\mathscr{M}_t$ has the form (6.2). The $(i,j)^{th}$ entry of the matrix M_t is

$$M_t^{i,j} = \mathscr{M}_t^{ij} = \sum_{k=1}^{d} \int_0^t D_s^k(X_t^i) D_s^k(X_t^j) ds = \sum_{k=1}^{d} \int_0^t D_s^k(X_t^i) D_s^k(Z_t^j) ds.$$

Therefore M_t is the Malliavin matrix corresponding to Z_t. □

Proof of Proposition 6.1.2. Note that $J_t^{n+1,i} = \frac{\partial}{\partial x^i}\zeta_t = 0$ for any $i \in \{1, \ldots, n\}$ therefore J_t has the form

$$J_t = \begin{pmatrix} \tilde{J}_t & a \\ 0 & b \end{pmatrix}$$

for some random real numbers a, b and a random $n \times n$ invertible matrix $\tilde{J}_t$. This implies that

$$J_t^{-1} = \begin{pmatrix} \tilde{J}_t^{-1} & -\tilde{J}_t^{-1} a b^{-1} \\ 0 & b^{-1} \end{pmatrix}$$

Now by Lemma 6.1.1 we have that

$$\begin{aligned}
\mathscr{C}_t &= J_t^{-1}\mathscr{M}_t(J_t^{-1})^T \\
&= \begin{pmatrix} \tilde{J}_t^{-1} & -\tilde{J}_t^{-1}ab^{-1} \\ 0 & b^{-1} \end{pmatrix} \begin{pmatrix} M_t & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} (\tilde{J}_t^{-1})^T & 0 \\ -\tilde{J}_t^{-1}ab^{-1} & b^{-1} \end{pmatrix} \\
&= \begin{pmatrix} \tilde{J}_t^{-1}M_t(\tilde{J}_t^{-1})^T & 0 \\ 0 & 0 \end{pmatrix}.
\end{aligned}$$

Let $C_t = \tilde{J}_t^{-1}M_t(\tilde{J}_t^{-1})^T$ then it remains to show that C_t is invertible.

It is sufficient to show that $\ker(C_t) = \{0\}$ almost surely, that is if there exists $v(\omega) \in \mathbb{R}^n$ such that $v^T C_t v = 0$ implies $v = 0$ almost surely. Note that

$$0 = v^T(\omega)C_t(\omega)v(\omega) = \sum_{k=1}^{d} \int_0^t |v^T J_t^{-1} V_k(X_s, Y_s)|^2 ds.$$

Therefore if $v(\omega) \in \ker(C_t(\omega))$ then v is orthogonal to the space

$$K_s := \operatorname{span}\{J_r^{-1}V_k(X_r) : 0 \leq r \leq s, k = 1, \ldots, d\}.$$

Hence it is sufficient to show that $K_s = \mathbb{R}^n$.

Note that the family of vector spaces $\{K_s : s \geq 0\}$ is increasing and set $K_{0+} := \cap_{s>0} K_s$. By the Blumenthal zero-one law, see Theorem 7.17 in [60], K_{0+} is a deterministic space with probability one. Define the stopping time

$$\tau = \inf\{s > 0 : \dim K_s > \dim K_{0+}\}.$$

Note that $\tau > 0$ with probability one. Let v be orthogonal to K_{0+} and non-zero, then we have $v \perp K_s$ if $s < \tau$, that is,

$$v^T J_t^{-1} V_k(X_s, Y_s) = 0, \quad k \in \{1, \ldots, d\}, s < \tau.$$

This follows since $K_{0+} \subseteq K_s$ for all $s > 0$ and for $s < \tau$ we have that $\dim(K_{0+}) = \dim(K_s)$.

Recall the set $\mathcal{R}_m$ was defined in (2.4), we shall denote by $\Delta_k(x)$ to be the vector space spanned by the vectors of $\mathcal{R}_k$ evaluated at the point x.

By following the proof of [35, Theorem 2.3.2] we obtain that v is orthogonal to $\Delta_k(x_0)$ and hence obtain that $\Delta_k(x_0) \subseteq K_{0+}$ for all $k \in \mathbb{N}$. By setting $k = m$ we have that $\mathbb{R}^n \subseteq K_{0+} \subseteq K_s \subseteq \mathbb{R}^n$. Therefore $\ker(C_t) = \{0\}$ and we have that C_t is invertible. $\square$

Bibliography

[1] T. Cass, D. Crisan, P. Dobson, and M. Ottobre, "Long-time behaviour of degenerate diffusions: UFG-type SDEs and time-inhomogeneous hypoelliptic processes," *arXiv preprint arXiv:1805.01350*, 2018.

[2] D. Crisan, P. Dobson, and M. Ottobre, "Uniform in time convergence for the Euler method and a pathwise approach to derivative estimates for diffusion semigroups," *Preprint*, 2019.

[3] L. Hörmander, "Hypoelliptic second order differential equations," *Acta Mathematica*, vol. 119, no. 1, pp. 147–171, 1967.

[4] L. R. Bellet, "Ergodic properties of markov processes," in *Open Quantum Systems II: The Markovian Approach* (S. Attal, A. Joye, and C.-A. Pillet, eds.), ch. 1, pp. 1–39, Berlin, Heidelberg: Springer, 2006.

[5] M. Hairer, "How hot can a heat bath get?," *Communications in Mathematical Physics*, vol. 292, no. 1, pp. 131–177, 2009.

[6] M. Ottobre and G. Pavliotis, "Asymptotic analysis for the generalized Langevin equation," *Nonlinearity*, vol. 24, no. 5, p. 1629, 2011.

[7] J.-P. Eckmann and M. Hairer, "Spectral properties of hypoelliptic operators," *Communications in mathematical physics*, vol. 235, no. 2, pp. 233–253, 2003.

[8] L. Lorenzi and M. Bertoldi, *Analytical Methods for Markov Semigroups*. Monographs and Research Notes in Mathematics, CRC Press, 2006.

[9] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*. Cambridge university press, 2014.

[10] G. Da Prato and J. Zabczyk, *Ergodicity for infinite dimensional systems*, vol. 229. Cambridge University Press, 1996.

[11] S. Cerrai, *Second order PDE's in finite and infinite dimension: a probabilistic approach*, vol. 1762. Springer Science & Business Media, 2001.

[12] J. L. Doob, *Stochastic processes*, vol. 101. New York Wiley, 1953.

[13] V. I. Bogachev, G. Da Prato, and M. Röckner, "Regularity of invariant measures for a class of perturbed ornstein-uhlenbeck operators," *Nonlinear Differential Equations and Applications NoDEA*, vol. 3, no. 2, pp. 261–268, 1996.

[14] V. I. Bogachev, G. Da Prato, and M. Röckner, "On parabolic equations for measures," *Communications in Partial Differential Equations*, vol. 33, no. 3, pp. 397–418, 2008.

[15] V. Bogachev, G. Da Prato, and M. Röckner, "Parabolic equations for measures on infinite-dimensional spaces," in *Doklady Mathematics*, vol. 78, pp. 544–549, Springer, 2008.

[16] V. I. Bogachev, G. Da Prato, M. Röckner, and W. Stannat, "Uniqueness of solutions to weak parabolic equations for measures," *Bulletin of the London Mathematical Society*, vol. 39, no. 4, pp. 631–640, 2007.

[17] V. I. Bogachev, G. Da Prato, and M. Röckner, "Existence of solutions to weak parabolic equations for measures," *Proceedings of the London Mathematical Society*, vol. 88, no. 3, pp. 753–774, 2004.

[18] R. Hermann, "On the accessibility problem in control theory," in *International Symposium on Nonlinear Differential Equations and Nonlinear Mechanics*, pp. 325–332, Elsevier, 1963.

[19] C. Lobry, "Controllability of nonlinear systems on compact manifolds," *SIAM Journal on Control*, vol. 12, no. 1, pp. 1–4, 1974.

[20] H. J. Sussmann, "Orbits of families of vector fields and integrability of distributions," *Transactions of the American Mathematical Society*, vol. 180, pp. 171–188, 1973.

[21] S. Kusuoka and D. Stroock, "Applications of the Malliavin calculus, part I," *North-Holland Mathematical Library*, vol. 32, pp. 271–306, 1984.

[22] S. Kusuoka and D. Stroock, "Applications of the Malliavin calculus. II," 1985.

[23] S. Kusuoka and D. Stroock, "Applications of the Malliavin calculus, part III," *Journal of The Faculty of Science, The University of Tokyo, Section IA, Mathematics*, vol. 34, no. 2, pp. 391–4421, 1987.

[24] S. Kusuoka, "Malliavin calculus revisited," *J. Math. Sci. Univ. Tokyo*, pp. 261–277, 2003.

[25] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and geometry of Markov diffusion operators*, vol. 348. Springer Science & Business Media, 2013.

[26] J. Barré, P. Dobson, M. Ottobre, and E. Zatorska, "Averaging for an infinite dimensional particle system," *work in progress*.

[27] P. Dobson and M. Ottobre, "Uniform in time averaging for sdes," *work in progress*.

[28] Ľ. Baňas, Z. Brzeźniak, M. Neklyudov, M. Ondreját, and A. Prohl, "Ergodicity for a stochastic geodesic equation in the tangent bundle of the 2D sphere," *Czechoslovak Mathematical Journal*, vol. 65, pp. 617–657, Sep 2015.

[29] D. Crisan, C. Litterer, and T. Lyons, "Kusuoka–Stroock gradient bounds for the solution of the filtering equation," *Journal of Functional Analysis*, vol. 268, no. 7, pp. 1928–1971, 2015.

[30] D. Crisan and F. Delarue, "Sharp derivative bounds for solutions of degenerate semi-linear partial differential equations," *Journal of Functional Analysis*, vol. 263, no. 10, pp. 3024–3101, 2012.

[31] D. Crisan, K. Manolarakis, and C. Nee, "Cubature methods and applications," *Lecture Notes in Mathematics*, vol. 2081, pp. 203–316, 2013.

[32] D. Crisan and M. Ottobre, "Pointwise gradient bounds for degenerate semigroups (of UFG type)," in *Proc. R. Soc. A*, vol. 472, p. 20160442, The Royal Society, 2016.

[33] M. Hairer, "On Malliavin's proof of Hörmander's theorem," *Bulletin des sciences mathematiques*, vol. 135, no. 6-7, pp. 650–666, 2011.

[34] D. Williams, *Stochastic Integrals: Proceedings of the LMS Durham Symposium, July 7-17, 1980.* Lecture Notes in Mathematics, Springer Berlin Heidelberg, 1981.

[35] D. Nualart, *The Malliavin calculus and related topics*, vol. 1995. Springer, 2006.

[36] J.-M. Bismut, "Martingales, the malliavin calculus and hypoellipticity under general hörmander's conditions," *Probability Theory and Related Fields*, vol. 56, no. 4, pp. 469–505, 1981.

[37] L. C. G. Rogers and D. Williams, *Diffusions, Markov processes and martingales: Volume 2, Itô calculus*, vol. 2. Cambridge university press, 2000.

[38] P. Cattiaux and L. Mesnager, "Hypoelliptic non-homogeneous diffusions," *Probability Theory and Related Fields*, vol. 123, no. 4, pp. 453–483, 2002.

[39] D. Crisan and E. McMurray, "Cubature on Wiener space for McKean–Vlasov SDEs with smooth scalar interaction," *The Annals of Applied Probability*, vol. 29, no. 1, pp. 130–177, 2019.

[40] M. Kunze, L. Lorenzi, and A. Lunardi, "Nonautonomous Kolmogorov parabolic equations with unbounded coefficients," *Transactions of the American Mathematical Society*, vol. 362, no. 1, pp. 169–198, 2010.

[41] G. Da Prato and M. Röckner, "A note on evolution systems of measures for time-dependent stochastic differential equations," in *Seminar on Stochastic Analysis, Random Fields and Applications V*, pp. 115–122, Springer, 2007.

[42] J. Schiltz, "Time depending Malliavin calculus on manifolds and application to nonlinear filtering," *Probability and Mathematical Statistics*, vol. 18, no. 2, pp. 319–334, 1998.

[43] P. Dobson, "A pathwise approach to the Bakry-Emery theory for derivative estimates for markov semigroups," *work in progress.*

[44] A. Lunardi, "On the Ornstein-Uhlenbeck operator in L^2 spaces with respect to invariant measures," *Transactions of the American Mathematical Society*, vol. 349, no. 1, pp. 155–169, 1997.

[45] M. Arnaudon and A. Thalmaier, "Bismut type differentiation of semigroups," *Probability Theory and Mathematical Statistics*, pp. 23–32, 1999.

[46] F. Dragoni, V. Kontis, and B. Zegarliński, "Ergodicity of markov semigroups with Hörmander type generators in infinite dimensions," *Potential Analysis*, pp. 1–29, 2012.

[47] M. Ottobre, "Asymptotic analysis for markovian models in non-equilibrium statistical mechanics," 2012.

[48] S. Kusuoka and D. Stroock, "Long time estimates for the heat kernel associated with a uniformly subelliptic symmetric second order operator," *Annals of mathematics*, vol. 127, no. 1, pp. 165–189, 1988.

[49] V. Kontis, M. Ottobre, and B. Zegarlinski, "Markov semigroups with hypocoercive-type generator in infinite dimensions: ergodicity and smoothing," *Journal of Functional Analysis*, vol. 270, no. 9, pp. 3173–3223, 2016.

[50] C. Nee, *Sharp Gradient Bounds for the Diffusion Semigroup*. Ph.d. thesis, 2011.

[51] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni, *Stratified Lie Groups and Potential Theory for their Sub-Laplacians*. Springer Monographs in Mathematics, Springer, 2007.

[52] M. D. Donsker and S. S. Varadhan, "Asymptotic evaluation of certain markov process expectations for large time, I," *Communications on Pure and Applied Mathematics*, vol. 28, no. 1, pp. 1–47, 1975.

[53] M. Donsker and S. Varadhan, "Asymptotic evaluation of certain markov process expectations for large time, II," *Communications on Pure and Applied Mathematics*, vol. 28, no. 2, pp. 279–301, 1975.

[54] M. Donsker and S. Varadhan, "Asymptotic evaluation of certain markov process expectations for large time, III," *Communications on pure and applied Mathematics*, vol. 29, no. 4, pp. 389–461, 1976.

[55] M. D. Donsker and S. S. Varadhan, "Asymptotic evaluation of certain markov process expectations for large time, IV," *Communications on Pure and Applied Mathematics*, vol. 36, no. 2, pp. 183–212, 1983.

[56] D. Crisan and S. Ghazali, "On the convergence rates of a general class of weak approximations of SDEs," in *Stochastic Differential Equations: Theory And Applications: A Volume in Honor of Professor Boris L Rozovskii*, pp. 221–248, World Scientific, 2007.

[57] E. D. Sontag, *Mathematical control theory: deterministic finite dimensional systems*, vol. 6. Springer Science & Business Media, 2013.

[58] A. Isidori, *Nonlinear control systems*. Springer Science & Business Media, 2013.

[59] A. Bellaiche and J. Risler, *Sub-Riemannian Geometry*. Progress in Mathematics, Birkhäuser Basel, 2012.

[60] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics, Springer New York, 1991.

[61] H. Kunita, "Stochastic differential equations and stochastic flows of diffeomorphisms," in *Ecole d'été de probabilités de Saint-Flour XII-1982*, pp. 143–303, Springer, 1984.

[62] L. Angiuli, L. Lorenzi, and A. Lunardi, "Hypercontractivity and asymptotic behavior in nonautonomous Kolmogorov equations," *Communications in Partial Differential Equations*, vol. 38, no. 12, pp. 2049–2080, 2013.

[63] M. Geissert and A. Lunardi, "Asymptotic behavior and hypercontractivity in non-autonomous Ornstein–Uhlenbeck equations," *Journal of the London Mathematical Society*, vol. 79, no. 1, pp. 85–106, 2009.

[64] P. Florchinger, "Malliavin calculus with time dependent coefficients and application to nonlinear filtering," *Probability theory and related fields*, vol. 86, no. 2, pp. 203–223, 1990.

[65] S. Taniguchi, "Malliavin's stochastic calculus of variations for manifold-valued Wiener functionals and its applications," *Probability Theory and Related Fields*, vol. 65, no. 2, pp. 269–290, 1983.

[66] S. Watanabe, "Malliavin's calculus in terms of generalized Wiener functionals," in *Theory and Application of Random Fields* (G. Kallianpur, ed.), (Berlin, Heidelberg), pp. 284–290, Springer Berlin Heidelberg, 1983.

[67] F. Rampazzo and H. J. Sussmann, "Commutators of flow maps of nonsmooth vector fields," *Journal of Differential Equations*, vol. 232, no. 1, pp. 134–175, 2007.

[68] W. Rudin, *Real and complex analysis.* Mathematics series, McGraw-Hill, 1987.

[69] J. M. Lee, "Smooth manifolds," in *Introduction to Smooth Manifolds*, pp. 1–31, Springer, 2013.

[70] T. W. Gamelin and R. E. Greene, *Introduction to topology.* Courier Corporation, 1999.